

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 19:01:11 ; Search time 503 Seconds
(without alignments)
236.522 Million cell updates/sec

Title: US-09-980-953-256

Perfect score: 20
Sequence: 1 cgtgtgtctgtgtagtccc 20

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	10	US-09-851-871-255
2	20	100.0	20	17	US-10-444-206-255
3	20	100.0	20	18	US-10-641-962-255
4	20	100.0	20	19	US-10-858-500-623
5	20	100.0	124	9	US-09-837-867A-39
6	20	100.0	124	10	US-09-962-969-39
7	20	100.0	124	18	US-10-643-768-39
8	20	100.0	598	14	US-09-796-692-7754
9	20	100.0	598	14	US-10-040-862-7754
10	20	100.0	598	17	US-10-057-475B-7754
11	20	100.0	598	17	US-10-154-884B-7754

Sequence 7754, Ap
Sequence 25, Appl
Sequence 25, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 22, Appl
Sequence 1, Appl
Sequence 296, App
Sequence 296, App
Sequence 297, App
Sequence 297, App
Sequence 7817, Ap
Sequence 7817, Ap
Sequence 7817, Ap
Sequence 7817, Ap
Sequence 7817, Ap
Sequence 120635,
Sequence 18097, A
Sequence 1063, Ap
Sequence 574, App
Sequence 613, App
Sequence 719033,
Sequence 23858, A
Sequence 1153, Ap
Sequence 290491,
Sequence 290491,
Sequence 290491,
Sequence 290491,
Sequence 146229,
Sequence 146229,
Sequence 146229,
Sequence 164600,
Sequence 164601,
Sequence 164601,
Sequence 14540, A
Sequence 340, Appl
Sequence 340, Appl
Sequence 645, App
Sequence 645, App
Sequence 645, App
Sequence 16, Appl
Sequence 324413,
Sequence 10145, A
Sequence 30483, A
Sequence 23283, A
Sequence 336, App
Sequence 29269, A
Sequence 29269, A
Sequence 320, App
Sequence 320, App
Sequence 9, Appl
Sequence 33553, A
Sequence 37, Appl
Sequence 35400, A
Sequence 12622, A
Sequence 48543, A
Sequence 48543, A
Sequence 12763, A
Sequence 216, App
Sequence 10001, A
Sequence 21972, A
Sequence 655, App
Sequence 11023, A
Sequence 3653, Ap
Sequence 966, Ap
Sequence 172479,
Sequence 23, Appl

85	15.2	76.0	598	13	US-10-027-632-194904	Sequence 194904,	C 158	14.8	74.0	564	17	US-10-027-632-267664	Sequence 267664,
86	15.2	76.0	598	13	US-10-027-632-194905	Sequence 194905,	C 159	14.8	74.0	564	17	US-10-027-632-267666	Sequence 267666,
87	15.2	76.0	598	13	US-10-027-632-194906	Sequence 194906,	C 160	14.8	74.0	575	13	US-10-027-632-269148	Sequence 269148,
88	15.2	76.0	598	17	US-10-027-632-194905	Sequence 194905,	C 161	14.8	74.0	575	13	US-10-027-632-269149	Sequence 269149,
89	15.2	76.0	603	10	US-09-814-353-16350	Sequence 16350, A	C 162	14.8	74.0	575	13	US-10-027-632-269150	Sequence 269150,
C 90	15.2	76.0	641	13	US-10-027-632-133093	Sequence 133093,	C 163	14.8	74.0	575	17	US-10-027-632-269148	Sequence 269148,
C 91	15.2	76.0	641	17	US-10-027-632-133093	Sequence 133093,	C 164	14.8	74.0	575	17	US-10-027-632-269149	Sequence 269149,
C 92	15.2	76.0	685	13	US-10-027-632-142147	Sequence 142147,	C 165	14.8	74.0	575	17	US-10-027-632-269150	Sequence 269150,
C 93	15.2	76.0	685	17	US-10-027-632-142147	Sequence 142147,	C 166	14.8	74.0	576	13	US-10-027-632-41732	Sequence 41732, A
C 94	15.2	76.0	993	17	US-10-424-599-23225	Sequence 23225, A	C 167	14.8	74.0	576	17	US-10-027-632-41732	Sequence 41732, A
C 95	15.2	76.0	1005	15	US-10-295-027-701	Sequence 701, App	C 168	14.8	74.0	585	13	US-10-027-632-63052	Sequence 63052, A
C 96	15.2	76.0	1070	13	US-10-098-841-168	Sequence 168, App	C 169	14.8	74.0	585	13	US-10-027-632-63938	Sequence 63938, A
C 97	15.2	76.0	2258	13	US-10-027-632-103137	Sequence 103137,	C 170	14.8	74.0	585	13	US-10-027-632-297454	Sequence 297454,
C 98	15.2	76.0	2258	17	US-10-027-632-103137	Sequence 103137,	C 171	14.8	74.0	585	17	US-10-027-632-63052	Sequence 63052, A
C 99	15.2	76.0	2264	17	US-10-421-554-71	Sequence 71, Appl	C 172	14.8	74.0	585	17	US-10-027-632-63938	Sequence 63938, A
C 100	15.2	76.0	3828	15	US-10-128-714-2125	Sequence 7125, App	C 173	14.8	74.0	585	17	US-10-027-632-297454	Sequence 297454,
C 101	15.2	76.0	3853	17	US-10-264-049-681	Sequence 681, App	C 174	14.8	74.0	601	16	US-10-417-510-21	Sequence 21, Appl
C 102	15.2	76.0	4079	15	US-10-128-714-1125	Sequence 1125, App	C 175	14.8	74.0	620	18	US-10-723-860-1656	Sequence 1656, App
C 103	15.2	76.0	4308	15	US-10-128-714-1125	Sequence 1125, App	C 176	14.8	74.0	631	13	US-10-027-632-79613	Sequence 79613, A
C 104	15.2	76.0	4627	15	US-10-128-714-6125	Sequence 6125, App	C 177	14.8	74.0	631	13	US-10-027-632-301275	Sequence 301275, A
C 105	15.2	76.0	6079	15	US-10-128-714-125	Sequence 125, App	C 178	14.8	74.0	631	17	US-10-027-632-79613	Sequence 79613, A
C 106	15.2	76.0	6089	18	US-10-758-672A-5	Sequence 5, Appl	C 179	14.8	74.0	631	17	US-10-027-632-301275	Sequence 301275, A
C 107	15.2	76.0	6627	13	US-10-128-714-5125	Sequence 5125, App	C 180	14.8	74.0	631	17	US-10-424-599-75120	Sequence 75120, A
C 108	15.2	76.0	30722	13	US-10-087-192-457	Sequence 457, App	C 181	14.8	74.0	654	17	US-10-424-599-14315	Sequence 14315, A
C 109	15.2	76.0	73725	17	US-10-052-482-196	Sequence 196, App	C 182	14.8	74.0	654	18	US-10-767-701-25770	Sequence 25770, A
C 110	15.2	76.0	80423	18	US-10-367-094-41	Sequence 41, Appl	C 183	14.8	74.0	705	13	US-10-027-632-267668	Sequence 267668,
C 111	15.2	76.0	120144	18	US-10-322-281-143	Sequence 143, App	C 184	14.8	74.0	705	16	US-10-417-510-3	Sequence 3, Appl
C 112	15.2	76.0	143068	9	US-09-967-768A-316	Sequence 316, App	C 185	14.8	74.0	705	17	US-10-027-632-267668	Sequence 267668,
C 113	15.2	76.0	309996	18	US-10-719-993-6785	Sequence 6785, App	C 186	14.8	74.0	728	13	US-10-027-632-19728	Sequence 19728, A
C 114	15.2	76.0	684973	9	US-09-263-959-1	Sequence 1, Appl	C 187	14.8	74.0	728	17	US-10-027-632-19728	Sequence 19728, A
C 115	15	75.0	624	18	US-10-437-963-63738	Sequence 63738, A	C 188	14.8	74.0	741	18	US-10-425-115-150893	Sequence 150893, A
C 116	15	75.0	843	18	US-10-437-963-69461	Sequence 69461, A	C 189	14.8	74.0	741	18	US-10-425-115-150893	Sequence 150893, A
C 117	15	75.0	1048	17	US-10-424-599-118922	Sequence 118922, A	C 190	14.8	74.0	741	18	US-10-425-115-150893	Sequence 150893, A
C 118	14.8	74.0	25	19	US-10-719-900-51824	Sequence 51824, A	C 191	14.8	74.0	932	17	US-10-425-114-4836	Sequence 4836, App
C 119	14.8	74.0	25	19	US-10-719-900-192211	Sequence 192211,	C 192	14.8	74.0	946	17	US-10-425-114-3646	Sequence 3646, App
C 120	14.8	74.0	25	19	US-10-719-900-908657	Sequence 908657,	C 193	14.8	74.0	955	18	US-10-684-422-17	Sequence 17, Appl
C 121	14.8	74.0	201	18	US-10-719-993-12788	Sequence 12788, A	C 194	14.8	74.0	957	19	US-10-774-355A-359	Sequence 359, App
C 122	14.8	74.0	339	18	US-10-767-701-16424	Sequence 16424, A	C 195	14.8	74.0	960	19	US-10-774-355A-408	Sequence 408, App
C 123	14.8	74.0	351	17	US-10-424-599-52826	Sequence 52826, A	C 196	14.8	74.0	991	18	US-10-425-115-35151	Sequence 35151, A
C 124	14.8	74.0	413	18	US-10-425-115-162973	Sequence 162973,	C 197	14.8	74.0	993	18	US-10-437-963-43376	Sequence 43376, A
C 125	14.8	74.0	464	13	US-10-027-632-311653	Sequence 311653,	C 198	14.8	74.0	1014	18	US-10-425-115-33505	Sequence 33505, A
C 126	14.8	74.0	464	13	US-10-027-632-311653	Sequence 311653,	C 199	14.8	74.0	1042	17	US-10-424-599-98223	Sequence 98223, A
C 127	14.8	74.0	464	17	US-10-027-632-311653	Sequence 311653,	C 200	14.8	74.0	1087	18	US-10-425-115-35152	Sequence 35152, A
C 128	14.8	74.0	464	17	US-10-027-632-311654	Sequence 311654,	C 201	14.8	74.0	1072	18	US-10-739-930-5110	Sequence 5110, App
C 129	14.8	74.0	465	17	US-10-424-599-137924	Sequence 137924, A	C 202	14.8	74.0	1311	17	US-10-369-493-47015	Sequence 47015, A
C 130	14.8	74.0	497	17	US-10-242-535A-24422	Sequence 24422, A	C 203	14.8	74.0	1565	17	US-10-425-114-27541	Sequence 27541, A
C 131	14.8	74.0	497	17	US-10-085-783A-24422	Sequence 24422, A	C 204	14.8	74.0	1815	18	US-10-425-115-175250	Sequence 175250, A
C 132	14.8	74.0	497	18	US-10-425-115-40003	Sequence 40003, A	C 205	14.8	74.0	1873	9	US-09-834-975-756	Sequence 756, App
C 133	14.8	74.0	512	13	US-10-027-632-267660	Sequence 267660,	C 206	14.8	74.0	1873	9	US-09-834-975-757	Sequence 757, App
C 134	14.8	74.0	512	13	US-10-027-632-267665	Sequence 267665,	C 207	14.8	74.0	2191	10	US-09-931-733-1	Sequence 1, Appl
C 135	14.8	74.0	512	13	US-10-027-632-267667	Sequence 267667,	C 208	14.8	74.0	2191	10	US-10-437-963-33304	Sequence 33304, A
C 136	14.8	74.0	512	17	US-10-027-632-267660	Sequence 267660,	C 209	14.8	74.0	3512	13	US-10-027-632-114278	Sequence 114278, A
C 137	14.8	74.0	512	17	US-10-027-632-267665	Sequence 267665,	C 210	14.8	74.0	3512	13	US-10-027-632-114279	Sequence 114279, A
C 138	14.8	74.0	512	17	US-10-027-632-267667	Sequence 267667,	C 211	14.8	74.0	3512	17	US-10-027-632-114278	Sequence 114278, A
C 139	14.8	74.0	519	9	US-09-918-995-5944	Sequence 5944, App	C 212	14.8	74.0	3512	9	US-10-027-632-114279	Sequence 114279, A
C 140	14.8	74.0	529	9	US-09-864-761-7787	Sequence 7787, App	C 213	14.8	74.0	4836	18	US-09-764-869-1735	Sequence 1735, App
C 141	14.8	74.0	538	18	US-10-767-701-18075	Sequence 18075, A	C 214	14.8	74.0	7537	17	US-10-091-504-1735	Sequence 1735, App
C 142	14.8	74.0	554	13	US-10-027-632-223235	Sequence 223235,	C 215	14.8	74.0	7537	17	US-10-027-577-1735	Sequence 1735, App
C 143	14.8	74.0	554	13	US-10-027-632-223236	Sequence 223236,	C 216	14.8	74.0	40304	11	US-09-997-722-280	Sequence 280, App
C 144	14.8	74.0	554	13	US-10-027-632-223237	Sequence 223237,	C 217	14.8	74.0	41322	15	US-10-024-396-13	Sequence 13, Appl
C 145	14.8	74.0	554	13	US-10-027-632-223238	Sequence 223238,	C 218	14.8	74.0	44990	17	US-10-052-482-217	Sequence 217, App
C 146	14.8	74.0	554	17	US-10-027-632-223235	Sequence 223235,	C 219	14.8	74.0	4836	9	US-09-927-602-38	Sequence 38, Appl
C 147	14.8	74.0	554	17	US-10-027-632-223236	Sequence 223236,	C 220	14.8	74.0	86574	18	US-10-775-169-173	Sequence 173, App
C 148	14.8	74.0	554	17	US-10-027-632-223237	Sequence 223237,	C 221	14.8	74.0	87869	18	US-10-717-597-194	Sequence 194, App
C 149	14.8	74.0	554	17	US-10-027-632-223238	Sequence 223238,	C 222	14.8	74.0	106938	18	US-10-322-281-566	Sequence 566, App
C 150	14.8	74.0	560	13	US-10-027-632-187475	Sequence 187475,	C 223	14.8	74.0	160552	18	US-10-697-628-11	Sequence 11, Appl
C 151	14.8	74.0	560	17	US-10-027-632-187476	Sequence 187476,	C 224	14.8	74.0	175338	13	US-10-087-192-1861	Sequence 1861, App
C 152	14.8	74.0	560	17	US-10-027-632-187475	Sequence 187475,	C 225	14.8	74.0	193303	15	US-10-081-327-37	Sequence 37, Appl
C 153	14.8	74.0	560	17	US-10-027-632-187476	Sequence 187476,	C 226	14.8	74.0	193303	15	US-10-081-327-37	Sequence 37, Appl
C 154	14.8	74.0	564	13	US-10-027-632-267661	Sequence 267661,	C 227	14.8	74.0	197997	9	US-09-822-246-3	Sequence 3, Appl
C 155	14.8	74.0	564	13	US-10-027-632-267664	Sequence 267664,	C 228	14.8	74.0	197997	18	US-10-469-028-3	Sequence 3, Appl
C 156	14.8	74.0	564	13	US-10-027-632-267666	Sequence 267666,	C 229	14.8	74.0	219352	18	US-10-322-281-45	Sequence 45, Appl
C 157	14.8	74.0	564	17	US-10-027-632-267661	Sequence 267661,	C 230	14.8	74.0	277616	18	US-10-367-094-83	Sequence 83, Appl

231	14.8	74.0	421609	18	US-10-367-094-122	Sequence 122, App	304	14.4	72.0	5253	14	US-10-080-943-3	Sequence 3, Appli
232	14.8	74.0	775062	18	US-10-719-993-6844	Sequence 6844, App	305	14.4	72.0	5549	18	US-10-723-860-5037	Sequence 5037, Ap
233	14.8	74.0	1223197	13	US-10-027-632-179264	Sequence 179264, A	306	14.4	72.0	8764	10	US-09-764-891-10019	Sequence 10019, A
234	14.8	74.0	1223197	17	US-10-027-632-179264	Sequence 179264, A	307	14.4	72.0	10445	10	US-09-764-891-6380	Sequence 6380, Ap
235	14.8	74.0	1691139	14	US-10-067-514-1	Sequence 1, Appli	308	14.4	72.0	14747	16	US-10-286-926-42	Sequence 42, Appl
236	14.8	74.0	1691139	17	US-10-419-723-1	Sequence 1, Appli	309	14.4	72.0	18340	16	US-10-240-376A-130	Sequence 130, App
237	14.4	72.0	129	17	US-10-424-599-69773	Sequence 69773, A	310	14.4	72.0	21704	15	US-10-017-161-709	Sequence 709, App
238	14.4	72.0	1210	18	US-10-674-124A-12661	Sequence 12661, A	311	14.4	72.0	21704	15	US-10-292-798-621	Sequence 621, App
239	14.4	72.0	232	17	US-10-424-599-18683	Sequence 18683, A	312	14.4	72.0	25346	18	US-10-601-700-15	Sequence 15, Appl
240	14.4	72.0	271	16	US-10-029-386-14471	Sequence 14471, A	313	14.4	72.0	32134	17	US-10-242-355-936	Sequence 936, App
241	14.4	72.0	296	18	US-10-723-860-3871	Sequence 3871, App	314	14.4	72.0	32192	17	US-10-242-355-937	Sequence 937, App
242	14.4	72.0	302	17	US-10-424-599-104449	Sequence 104449, A	315	14.4	72.0	35460	18	US-10-322-696-94	Sequence 94, Appl
243	14.4	72.0	315	18	US-10-357-930-6846	Sequence 6846, App	316	14.4	72.0	96276	18	US-10-723-860-451	Sequence 451, App
244	14.4	72.0	328	18	US-10-357-930-6547	Sequence 6547, App	317	14.4	72.0	96596	17	US-10-052-482-10	Sequence 10, Appl
245	14.4	72.0	364	18	US-10-674-124A-21095	Sequence 21095, A	318	14.4	72.0	168198	18	US-10-322-696-98	Sequence 88, Appl
246	14.4	72.0	366	17	US-10-085-783A-23201	Sequence 23201, A	319	14.4	72.0	173808	13	US-10-003-806-10	Sequence 10, Appl
247	14.4	72.0	366	17	US-10-085-783A-23201	Sequence 23201, A	320	14.4	72.0	176080	17	US-10-235-132A-43	Sequence 43, Appl
248	14.4	72.0	369	13	US-10-027-632-50282	Sequence 50282, A	321	14.4	72.0	201143	17	US-10-240-425-1099	Sequence 1099, Ap
249	14.4	72.0	369	17	US-10-027-632-50282	Sequence 50282, A	322	14.4	72.0	247544	15	US-10-322-696-55	Sequence 55, Appl
250	14.4	72.0	394	17	US-10-424-599-5269	Sequence 5269, App	323	14.4	72.0	268685	15	US-10-052-071-22	Sequence 22, Appl
251	14.4	72.0	404	18	US-10-723-860-1273	Sequence 1273, App	324	14.4	72.0	268685	15	US-10-025-966A-22	Sequence 22, Appl
252	14.4	72.0	416	9	US-09-770-423-464	Sequence 464, App	325	14.4	72.0	310122	18	US-10-417-375-1	Sequence 1, Appli
253	14.4	72.0	420	18	US-10-357-930-36525	Sequence 36525, A	326	14.4	72.0	561515	18	US-10-741-601-5682	Sequence 5682, Ap
254	14.4	72.0	443	17	US-10-242-535A-7168	Sequence 7168, App	327	14.4	72.0	561515	19	US-10-741-600-17730	Sequence 17730, A
255	14.4	72.0	443	17	US-10-085-783A-7168	Sequence 7168, App	328	14.2	71.0	25	15	US-10-098-263B-7945	Sequence 7945, Ap
256	14.4	72.0	445	9	US-09-960-352-8746	Sequence 8746, App	329	14.2	71.0	25	19	US-10-719-900-332444	Sequence 332444, A
257	14.4	72.0	470	10	US-09-918-995-28591	Sequence 28591, App	330	14.2	71.0	25	19	US-10-719-900-528634	Sequence 528634, A
258	14.4	72.0	470	10	US-10-357-930-36818	Sequence 36818, A	331	14.2	71.0	25	19	US-10-719-900-588448	Sequence 588448, A
259	14.4	72.0	490	10	US-09-918-995-903	Sequence 903, App	332	14.2	71.0	25	19	US-10-719-900-701534	Sequence 701534, A
260	14.4	72.0	512	13	US-10-027-632-267663	Sequence 267663, A	333	14.2	71.0	25	19	US-10-719-900-719032	Sequence 719032, A
261	14.4	72.0	512	17	US-10-027-632-267663	Sequence 267663, A	334	14.2	71.0	50	17	US-10-131-827-181	Sequence 181, App
262	14.4	72.0	520	18	US-10-357-930-6546	Sequence 6546, App	335	14.2	71.0	54	17	US-10-440-850-2263	Sequence 2263, App
263	14.4	72.0	524	16	US-10-425-115-53758	Sequence 53758, A	336	14.2	71.0	75	18	US-10-674-124A-22635	Sequence 22635, A
264	14.4	72.0	525	16	US-10-029-386-12883	Sequence 12883, A	337	14.2	71.0	93	18	US-10-674-124A-4176	Sequence 4176, Ap
265	14.4	72.0	543	13	US-10-027-632-90425	Sequence 90425, A	338	14.2	71.0	95	9	US-09-864-761-22011	Sequence 22011, Ap
266	14.4	72.0	543	17	US-10-027-632-90425	Sequence 90425, A	339	14.2	71.0	111	18	US-10-674-124A-18235	Sequence 18235, A
267	14.4	72.0	545	9	US-09-925-300-266	Sequence 266, App	340	14.2	71.0	129	18	US-10-674-124A-18749	Sequence 18749, A
268	14.4	72.0	547	16	US-10-029-386-768	Sequence 768, App	341	14.2	71.0	138	18	US-10-674-124A-19715	Sequence 19715, A
269	14.4	72.0	551	13	US-10-027-632-51580	Sequence 51580, A	342	14.2	71.0	138	18	US-10-674-124A-19715	Sequence 19715, A
270	14.4	72.0	551	13	US-10-027-632-53839	Sequence 53839, A	343	14.2	71.0	139	18	US-10-767-701-19286	Sequence 19286, A
271	14.4	72.0	551	13	US-10-027-632-64643	Sequence 64643, A	344	14.2	71.0	158	18	US-10-674-124A-18552	Sequence 18552, A
272	14.4	72.0	551	13	US-10-027-632-321516	Sequence 321516, A	345	14.2	71.0	160	18	US-10-674-124A-909	Sequence 909, App
273	14.4	72.0	551	17	US-10-027-632-51580	Sequence 51580, A	346	14.2	71.0	165	18	US-10-674-124A-910	Sequence 910, App
274	14.4	72.0	551	17	US-10-027-632-53839	Sequence 53839, A	347	14.2	71.0	169	18	US-10-425-115-6203	Sequence 6203, App
275	14.4	72.0	551	17	US-10-027-632-64643	Sequence 64643, A	348	14.2	71.0	177	17	US-10-437-963-77372	Sequence 77372, A
276	14.4	72.0	551	17	US-10-027-632-321516	Sequence 321516, A	349	14.2	71.0	201	18	US-10-741-601-7589	Sequence 7589, App
277	14.4	72.0	564	13	US-10-027-632-267662	Sequence 267662, A	350	14.2	71.0	201	18	US-10-741-601-7665	Sequence 7665, App
278	14.4	72.0	564	17	US-10-027-632-267662	Sequence 267662, A	351	14.2	71.0	201	18	US-10-741-601-7677	Sequence 7677, App
279	14.4	72.0	882	18	US-10-357-930-21137	Sequence 21137, A	352	14.2	71.0	201	18	US-10-741-601-7677	Sequence 7677, App
280	14.4	72.0	882	18	US-10-357-930-26976	Sequence 26976, A	353	14.2	71.0	201	18	US-10-719-993-3608	Sequence 3608, App
281	14.4	72.0	933	19	US-10-774-355A-141	Sequence 141, App	354	14.2	71.0	201	18	US-10-719-993-37812	Sequence 37812, A
282	14.4	72.0	1139	17	US-10-424-599-43911	Sequence 43911, A	355	14.2	71.0	201	18	US-10-719-993-29153	Sequence 29153, A
283	14.4	72.0	1237	10	US-09-984-276-66	Sequence 66, Appl	356	14.2	71.0	201	18	US-10-719-993-29704	Sequence 29704, A
284	14.4	72.0	1237	10	US-09-984-271-66	Sequence 66, Appl	357	14.2	71.0	201	18	US-10-719-993-42824	Sequence 42824, A
285	14.4	72.0	1340	18	US-10-357-930-24176	Sequence 24176, A	358	14.2	71.0	201	19	US-10-741-600-21382	Sequence 21382, A
286	14.4	72.0	1340	18	US-10-357-930-25496	Sequence 25496, A	359	14.2	71.0	201	19	US-10-741-600-21458	Sequence 21458, A
287	14.4	72.0	1412	18	US-10-357-930-24391	Sequence 24391, A	360	14.2	71.0	201	19	US-10-741-600-21460	Sequence 21460, A
288	14.4	72.0	1412	18	US-10-357-930-24486	Sequence 24486, A	361	14.2	71.0	201	19	US-10-741-600-21470	Sequence 21470, A
289	14.4	72.0	1740	17	US-10-108-260A-1833	Sequence 1833, App	362	14.2	71.0	216	10	US-09-918-995-32728	Sequence 32728, A
290	14.4	72.0	2043	9	US-09-974-300-1653	Sequence 1653, App	363	14.2	71.0	223	9	US-09-864-761-18011	Sequence 18011, A
291	14.4	72.0	2254	13	US-09-925-301-453	Sequence 453, App	364	14.2	71.0	223	18	US-10-425-115-24218	Sequence 24218, A
292	14.4	72.0	2393	13	US-10-098-841-47	Sequence 47, Appl	365	14.2	71.0	238	9	US-10-674-124A-12659	Sequence 12659, A
293	14.4	72.0	2483	17	US-10-104-047-252	Sequence 252, App	366	14.2	71.0	245	9	US-09-922-217-576	Sequence 576, App
294	14.4	72.0	2520	17	US-10-159-563-107	Sequence 107, App	367	14.2	71.0	245	9	US-09-833-263-576	Sequence 576, App
295	14.4	72.0	2530	18	US-10-723-860-5768	Sequence 5768, App	368	14.2	71.0	245	13	US-10-025-380-576	Sequence 576, App
296	14.4	72.0	2725	9	US-09-964-824A-243	Sequence 243, App	369	14.2	71.0	259	17	US-10-424-599-110379	Sequence 110379, A
297	14.4	72.0	2725	9	US-09-880-107-2181	Sequence 2181, App	370	14.2	71.0	279	18	US-10-425-115-11452	Sequence 11452, A
298	14.4	72.0	2861	18	US-10-437-963-36373	Sequence 36373, A	371	14.2	71.0	314	10	US-09-918-995-19016	Sequence 19016, A
299	14.4	72.0	3820	18	US-10-723-860-5474	Sequence 5474, App	372	14.2	71.0	321	18	US-10-425-115-20683	Sequence 20683, A
300	14.4	72.0	3846	14	US-10-198-846-10269	Sequence 10269, A	373	14.2	71.0	328	18	US-10-425-115-32214	Sequence 32214, A
301	14.4	72.0	4775	17	US-10-191-803-88	Sequence 88, Appl	374	14.2	71.0	335	9	US-09-864-761-21277	Sequence 21277, A
302	14.4	72.0	4775	17	US-10-152-319A-1657	Sequence 1657, App	375	14.2	71.0	340	18	US-10-674-124A-2741	Sequence 2741, Ap
303	14.4	72.0	5253	9	US-09-880-107-1655	Sequence 1655, App	376	14.2	71.0	345	17	US-10-242-535A-38239	Sequence 38239, A

C 377	14.2	71.0	345	17	US-10-085-783A-38239	Sequence 38239, A	C 450	14.2	71.0	663	17	US-10-027-632-234171	Sequence 234171, A
C 378	14.2	71.0	348	17	US-10-242-535A-11103	Sequence 11103, A	C 451	14.2	71.0	668	13	US-10-027-632-114166	Sequence 114166, A
C 379	14.2	71.0	348	17	US-10-085-783A-11103	Sequence 11103, A	C 452	14.2	71.0	688	13	US-10-027-632-114167	Sequence 114167, A
C 380	14.2	71.0	358	9	US-09-864-761-4530	Sequence 4530, Ap	C 453	14.2	71.0	688	17	US-10-027-632-114166	Sequence 114166, A
C 381	14.2	71.0	362	18	US-10-437-963-57653	Sequence 57653, A	C 454	14.2	71.0	688	17	US-10-027-632-114167	Sequence 114167, A
C 382	14.2	71.0	365	18	US-10-674-124A-14420	Sequence 14420, A	C 455	14.2	71.0	690	13	US-10-027-632-240712	Sequence 240712, A
C 383	14.2	71.0	367	18	US-10-425-115-123687	Sequence 123687, A	C 456	14.2	71.0	690	17	US-10-027-632-240712	Sequence 240712, A
C 384	14.2	71.0	389	18	US-10-674-124A-4515	Sequence 4515, Ap	C 457	14.2	71.0	707	14	US-10-175-523-73	Sequence 73, Appl
C 385	14.2	71.0	389	18	US-10-674-124A-25377	Sequence 25377, A	C 458	14.2	71.0	712	13	US-10-027-632-167143	Sequence 167143, A
C 386	14.2	71.0	396	17	US-10-242-535A-10014	Sequence 10014, A	C 459	14.2	71.0	712	17	US-10-027-632-167143	Sequence 167143, A
C 387	14.2	71.0	396	17	US-10-242-535A-39623	Sequence 39623, A	C 460	14.2	71.0	713	13	US-10-027-632-150816	Sequence 150816, A
C 388	14.2	71.0	396	17	US-10-085-783A-10014	Sequence 10014, A	C 461	14.2	71.0	713	17	US-10-027-632-150816	Sequence 150816, A
C 389	14.2	71.0	396	17	US-10-085-783A-39623	Sequence 39623, A	C 462	14.2	71.0	713	18	US-10-357-930-45655	Sequence 45655, A
C 390	14.2	71.0	398	13	US-10-027-632-96085	Sequence 96085, A	C 463	14.2	71.0	713	18	US-10-357-930-45892	Sequence 45892, A
C 391	14.2	71.0	398	13	US-10-027-632-96085	Sequence 96085, A	C 464	14.2	71.0	721	13	US-10-027-632-19774	Sequence 19774, A
C 392	14.2	71.0	403	14	US-10-198-846-3149	Sequence 3149, Ap	C 465	14.2	71.0	721	17	US-10-027-632-19774	Sequence 19774, A
C 393	14.2	71.0	408	9	US-09-867-701-5113	Sequence 5113, Ap	C 466	14.2	71.0	734	17	US-10-287-092-21	Sequence 21, Appl
C 394	14.2	71.0	423	9	US-09-864-761-5233	Sequence 5233, Ap	C 467	14.2	71.0	738	18	US-10-425-115-110024	Sequence 110024, A
C 395	14.2	71.0	432	17	US-10-424-599-68806	Sequence 68806, A	C 468	14.2	71.0	739	17	US-10-287-092-23	Sequence 23, Appl
C 396	14.2	71.0	432	17	US-09-954-456-33	Sequence 33, Appl	C 469	14.2	71.0	744	17	US-10-425-114-17023	Sequence 17023, A
C 397	14.2	71.0	436	9	US-09-954-456-672	Sequence 672, App	C 470	14.2	71.0	747	13	US-10-027-632-148613	Sequence 148613, A
C 398	14.2	71.0	441	18	US-10-357-930-16066	Sequence 16066, A	C 471	14.2	71.0	747	17	US-10-027-632-148613	Sequence 148613, A
C 399	14.2	71.0	455	17	US-10-424-599-124659	Sequence 124659, A	C 472	14.2	71.0	748	13	US-10-027-632-161496	Sequence 161496, A
C 400	14.2	71.0	463	10	US-09-918-995-22038	Sequence 22038, A	C 473	14.2	71.0	748	17	US-10-027-632-161496	Sequence 161496, A
C 401	14.2	71.0	464	18	US-10-425-115-160314	Sequence 160314, A	C 474	14.2	71.0	767	13	US-10-027-632-154445	Sequence 154445, A
C 402	14.2	71.0	467	9	US-09-864-761-1250	Sequence 1250, Ap	C 475	14.2	71.0	767	17	US-10-027-632-154445	Sequence 154445, A
C 403	14.2	71.0	467	17	US-10-424-599-111248	Sequence 111248, A	C 476	14.2	71.0	774	9	US-09-731-872-134	Sequence 134, App
C 404	14.2	71.0	469	18	US-10-674-124A-21949	Sequence 21949, A	C 477	14.2	71.0	774	10	US-09-876-997-134	Sequence 134, App
C 405	14.2	71.0	470	9	US-09-864-761-6577	Sequence 6577, Ap	C 478	14.2	71.0	776	13	US-10-027-632-167142	Sequence 167142, A
C 406	14.2	71.0	471	9	US-09-960-352-3869	Sequence 3869, Ap	C 479	14.2	71.0	776	17	US-10-027-632-167142	Sequence 167142, A
C 407	14.2	71.0	473	11	US-09-969-034-847	Sequence 847, App	C 480	14.2	71.0	788	17	US-10-425-114-18128	Sequence 18128, A
C 408	14.2	71.0	478	10	US-09-918-995-2758	Sequence 2758, App	C 481	14.2	71.0	801	13	US-10-027-632-127379	Sequence 127379, A
C 409	14.2	71.0	483	18	US-10-437-963-78928	Sequence 78928, A	C 482	14.2	71.0	801	17	US-10-027-632-127379	Sequence 127379, A
C 410	14.2	71.0	493	13	US-10-027-632-283703	Sequence 283703, A	C 483	14.2	71.0	807	13	US-10-027-632-165907	Sequence 165907, A
C 411	14.2	71.0	493	17	US-10-027-632-283703	Sequence 283703, A	C 484	14.2	71.0	807	13	US-10-027-632-165907	Sequence 165907, A
C 412	14.2	71.0	495	10	US-09-918-995-37888	Sequence 37888, A	C 485	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 413	14.2	71.0	495	10	US-09-918-995-21557	Sequence 21557, A	C 486	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 414	14.2	71.0	520	18	US-10-425-115-113668	Sequence 113668, A	C 487	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 415	14.2	71.0	524	18	US-10-357-930-15827	Sequence 15827, A	C 488	14.2	71.0	807	17	US-10-027-632-165908	Sequence 165908, A
C 416	14.2	71.0	532	13	US-10-027-632-282909	Sequence 282909, A	C 489	14.2	71.0	815	13	US-10-027-632-127378	Sequence 127378, A
C 417	14.2	71.0	532	17	US-10-027-632-282909	Sequence 282909, A	C 490	14.2	71.0	815	17	US-10-027-632-164811	Sequence 164811, A
C 418	14.2	71.0	537	17	US-10-424-599-46076	Sequence 46076, A	C 491	14.2	71.0	819	13	US-10-027-632-164811	Sequence 164811, A
C 419	14.2	71.0	553	19	US-10-643-775-466	Sequence 466, App	C 492	14.2	71.0	823	18	US-10-425-115-176861	Sequence 176861, A
C 420	14.2	71.0	556	18	US-10-357-930-30160	Sequence 30160, A	C 493	14.2	71.0	823	18	US-10-425-115-176861	Sequence 176861, A
C 421	14.2	71.0	556	18	US-10-357-930-30344	Sequence 30344, A	C 494	14.2	71.0	824	17	US-10-425-114-18356	Sequence 18356, A
C 422	14.2	71.0	575	18	US-10-357-930-50074	Sequence 50074, A	C 495	14.2	71.0	826	13	US-10-027-632-155117	Sequence 155117, A
C 423	14.2	71.0	577	18	US-10-357-930-50142	Sequence 50142, A	C 496	14.2	71.0	826	17	US-10-027-632-155117	Sequence 155117, A
C 424	14.2	71.0	592	14	US-10-198-846-7919	Sequence 7919, Ap	C 497	14.2	71.0	836	13	US-10-027-632-159588	Sequence 159588, A
C 425	14.2	71.0	599	17	US-10-424-599-2472	Sequence 2472, Ap	C 498	14.2	71.0	836	13	US-10-027-632-159589	Sequence 159589, A
C 426	14.2	71.0	601	14	US-10-198-846-7978	Sequence 7978, Ap	C 499	14.2	71.0	836	13	US-10-027-632-159589	Sequence 159589, A
C 427	14.2	71.0	611	19	US-10-487-556-40	Sequence 40, Appl	C 500	14.2	71.0	836	17	US-10-027-632-159588	Sequence 159588, A
C 428	14.2	71.0	612	13	US-10-027-632-187176	Sequence 187176, A							
C 429	14.2	71.0	612	17	US-10-027-632-187176	Sequence 187176, A							
C 430	14.2	71.0	618	13	US-10-027-632-149994	Sequence 149994, A							
C 431	14.2	71.0	618	14	US-10-198-846-9013	Sequence 9013, Ap							
C 432	14.2	71.0	618	17	US-10-027-632-149994	Sequence 149994, A							
C 433	14.2	71.0	624	17	US-10-369-493-43481	Sequence 43481, A							
C 434	14.2	71.0	628	13	US-10-027-632-240010	Sequence 240010, A							
C 435	14.2	71.0	628	17	US-10-027-632-240010	Sequence 240010, A							
C 436	14.2	71.0	635	18	US-10-619-739-2051	Sequence 2051, Ap							
C 437	14.2	71.0	635	18	US-10-425-115-57413	Sequence 57413, A							
C 438	14.2	71.0	635	18	US-10-425-115-57413	Sequence 57413, A							
C 439	14.2	71.0	639	13	US-10-027-632-318713	Sequence 318713, A							
C 440	14.2	71.0	639	13	US-10-027-632-318713	Sequence 318713, A							
C 441	14.2	71.0	639	17	US-10-027-632-95409	Sequence 95409, A							
C 442	14.2	71.0	639	17	US-10-027-632-318713	Sequence 318713, A							
C 443	14.2	71.0	643	14	US-10-027-632-318713	Sequence 318713, A							
C 444	14.2	71.0	654	18	US-10-198-846-2100	Sequence 2100, Ap							
C 445	14.2	71.0	654	18	US-10-767-701-10777	Sequence 10777, A							
C 446	14.2	71.0	660	13	US-10-027-632-18941	Sequence 18941, A							
C 447	14.2	71.0	660	17	US-10-027-632-18941	Sequence 18941, A							
C 448	14.2	71.0	663	13	US-10-027-632-234170	Sequence 234170, A							
C 449	14.2	71.0	663	13	US-10-027-632-234171	Sequence 234171, A							
C 450	14.2	71.0	663	17	US-10-027-632-234170	Sequence 234170, A							

ALIGNMENTS

RESULT 1

US-09-851-871-255
; Sequence 255, Application US/09851871
; Publication No. US20030176374A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Clarence Frank
; APPLICANT: Karras, James G.
; APPLICANT: Vickers, Timothy A.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE: ISPH-0543
; CURRENT APPLICATION NUMBER: US/09/851,871
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/14471
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/326,186
; PRIOR FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 08/777,266
; PRIOR FILING DATE: 1996-12-31
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-851-871-255

Query Match 100.0%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||
Db 1 CGTGTCTCTGTCTAGTCCC 20

RESULT 2
US-10-444-206-255
; Sequence 255, Application US/10444206
; Publication No. US20040023917A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Clarence Frank
; APPLICANT: Vickers, Timothy A.
; APPLICANT: Karras, James G.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/444,206
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: 09/851,871
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/14471
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/326,186
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 08/777,266
; PRIOR FILING DATE: 1996-12-31
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-444-206-255

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||
Db 1 CGTGTCTCTGTCTAGTCCC 20

RESULT 3
US-10-641-962-255
; Sequence 255, Application US/10641962
; Publication No. US20040235164A1
; GENERAL INFORMATION:
; APPLICANT: Bennett et al.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE: 30566/39578
; CURRENT APPLICATION NUMBER: US/10/641,962
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 444

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic primer
US-10-641-962-255

Query Match 100.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||
Db 1 CGTGTCTCTGTCTAGTCCC 20

RESULT 4
US-10-858-500-623
; Sequence 623, Application US/10858500
; Publication No. US20050014257A1
; GENERAL INFORMATION:
; APPLICANT: Rosanne M. Crooke
; APPLICANT: Mark J. Graham
; TITLE OF INVENTION: MODULATION OF C-REACTIVE PROTEIN EXPRESSION
; FILE REFERENCE: BIOL0014US
; CURRENT APPLICATION NUMBER: US/10/858,500
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 09/912,724
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/475,272
; PRIOR FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/540,042
; PRIOR FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 627
; SEQ ID NO 623
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-858-500-623

Query Match 100.0%; Score 20; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||
Db 1 CGTGTCTCTGTCTAGTCCC 20

RESULT 5
US-09-837-867A-39/c
; Sequence 39, Application US/09837867A
; Patent No. US20020098542A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. US20020098542A1e1 Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-09-837-867A-39

Query Match      100.0%; Score 20; DB 9; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 6
US-09-962-969-39/c
; Sequence 39, Application US/09962969
; Publication No. US20030045703A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon
; APPLICANT: Nadler, Lee
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory
; FILE REFERENCE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,969
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/702,525
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: BWI-120CPUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-962-969-39

Query Match      100.0%; Score 20; DB 10; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38
```

```
RESULT 7
US-10-643-768-39/c
; Sequence 39, Application US/10643768
; Publication No. US20040192899A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/10/643,768
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/837,867
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-10-643-768-39

Query Match      100.0%; Score 20; DB 18; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 8
US-09-796-692-7754/c
; Sequence 7754, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
```

; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (574)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7754

Query Match 100.0%; Score 20; DB 9; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQTGTGCTGTGCTAGTCC 20
DB 63 CQTGTGCTGTGCTAGTCC 44

RESULT 9

US-10-040-862-7754/c
; Sequence 7754, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (574)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-7754

Query Match 100.0%; Score 20; DB 14; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQTGTGCTGTGCTAGTCC 20
DB 63 CQTGTGCTGTGCTAGTCC 44

RESULT 10

US-10-057-475B-7754/c
; Sequence 7754, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(598)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-7754

Query Match 100.0%; Score 20; DB 17; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQTGTGCTGTGCTAGTCC 20
DB 63 CQTGTGCTGTGCTAGTCC 44

```
RESULT 11
US-10-154-884B-7754/c
; Sequence 7754, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11250
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(598)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7754

Query Match 100.0%; Score 20; DB 17; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 63 CGTGTGCTGTGCTAGTCCC 44

RESULT 12
US-10-764-324-7754/c
; Sequence 7754, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11250
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7754
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(598)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7754

Query Match 100.0%; Score 20; DB 17; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 63 CGTGTGCTGTGCTAGTCCC 44

RESULT 13
US-09-441-411-25/c
; Sequence 25, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-25

Query Match 100.0%; Score 20; DB 10; Length 1112;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38
```

RESULT 14
US-10-762-128-25/c
; Sequence 25, Application US/10762128
; Publication No. US20040219161A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409C1
; CURRENT APPLICATION NUMBER: US/10/762,128
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 09/441,411
; FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-762-128-25

Query Match 100.0%; Score 20; DB 18; Length 1112;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15
US-08-592-711-3/c
; Sequence 3, Application US/08592711
; Publication No. US20020115214A1
; GENERAL INFORMATION:
; APPLICANT: June, Carl H.
; APPLICANT: Thompson, Craig B.
; APPLICANT: Nabel, Gary J.
; APPLICANT: Gray, Gary S.
; APPLICANT: Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,711
; FILING DATE: 26-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,816
; FILING DATE: 4-MAY-1995
; APPLICATION NUMBER: US 08/403,253
; FILING DATE: 10-MARCH-1995
; APPLICATION NUMBER: US 08/253,964
; FILING DATE: 3-JUNE-1994
; APPLICATION NUMBER: US 08/073,223
; FILING DATE: 4-JUNE-1993
; APPLICATION NUMBER: US 08/200,947
; FILING DATE: 23-FEB-1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7-APR-1992
; APPLICATION NUMBER: US 08/247,505

; FILING DATE: 23-MAY-1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7-APR-1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7-APR-1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16-JUNE-1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23-NOV-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-592-711-3

Query Match 100.0%; Score 20; DB 8; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 16
US-09-183-055-3/c
; Sequence 3, Application US/09183055
; Publication No. US20020076407A1
; GENERAL INFORMATION:
; APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
; APPLICANT: Gray, Gary S., Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating
; Proliferation Of T-Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/183,055
; FILING DATE: 30-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,253
; FILING DATE: March 10, 1995
; APPLICATION NUMBER: US 08/253,964
; FILING DATE: 3 JUNE 1994
; APPLICATION NUMBER: US 08/073,223
; FILING DATE: 4 JUNE 1993
; APPLICATION NUMBER: US 08/200,947

```
;
; FILING DATE: 23 FEB 1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/247,505
; FILING DATE: 23 MAY 1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25 MAR 1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16 JUNE 1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23 NOV 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Superko, Colleen
; REGISTRATION NUMBER: 39,850
; REFERENCE/DOCKET NUMBER: 36119-125US8
; TELEPHONE: (617) 526-6564
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-183-055-3
;
; Query Match 100.0%; Score 20; DB 9; Length 1120;
; Best Local Similarity 100.0%; Pred. No. 3.3;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CGTGTGCTGTGCTAGTCCC 20
; Db 57 CGTGTGCTGTGCTAGTCCC 38
;
; RESULT 17
; US-09-425-762-1/c
; Sequence 1, Application US/09425762
; Publication No. US20020086414A1
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 6605279e1 CTLA4/CD28 Ligands and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,762
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
;
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; US-09-425-762-1
;
; Query Match 100.0%; Score 20; DB 9; Length 1120;
; Best Local Similarity 100.0%; Pred. No. 3.3;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CGTGTGCTGTGCTAGTCCC 20
; Db 57 CGTGTGCTGTGCTAGTCCC 38
;
; RESULT 18
; US-09-837-867A-22/c
; Sequence 22, Application US/09837867A
; Patent No. US20020098542A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. US20020098542A1el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(1093)
; US-09-837-867A-22
;
; Query Match 100.0%; Score 20; DB 9; Length 1120;
; Best Local Similarity 100.0%; Pred. No. 3.3;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CGTGTGCTGTGCTAGTCCC 20
; Db 57 CGTGTGCTGTGCTAGTCCC 38
;
; RESULT 19
; US-09-962-969-22/c
; Sequence 22, Application US/09962969
; Publication No. US20030045703A1
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon
```

;/ Nadler, Lee
; TITLE OF INVENTION: No. US20030045703a1el Forms of T Cell Costimulatory
; Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,969
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/702,525
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: BWI-120CPUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-962-969-22

Query Match 100.0%; Score 20; DB 10; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
Db 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 20
US-350-202-3/c
; Sequence 3, Application US/09350202
; Publication No. US20030099643A1
; GENERAL INFORMATION:
; APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
; APPLICANT: Gray, Gary S., Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

;/ APPLICATION NUMBER: US/09/350,202
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/403,253
; FILING DATE: March 10, 1995
; APPLICATION NUMBER: US 08/253,964
; FILING DATE: 3 JUNE 1994
; APPLICATION NUMBER: US 08/073,223
; FILING DATE: 4 JUNE 1993
; APPLICATION NUMBER: US 08/200,947
; FILING DATE: 23 FEB 1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/247,505
; FILING DATE: 23 MAY 1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25 MAR 1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16 JUNE 1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23 NOV 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-09-350-202-3

Query Match 100.0%; Score 20; DB 10; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
Db 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 21
US-10-390-330-3/c
; Sequence 3, Application US/10390330
; Publication No. US20040001829A1
; GENERAL INFORMATION:
; APPLICANT: June, Carl H.
; APPLICANT: Thompson, Craig B.
; APPLICANT: Nabel, Gary J.
; APPLICANT: Gray, Gary S.
; APPLICANT: Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/390,330
FILING DATE: March 17, 2003
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/435,816A
FILING DATE: May 4, 1995
CLASSIFICATION:
APPLICATION NUMBER: US 08/403,253
FILING DATE: 10 MARCH 1995
APPLICATION NUMBER: US 08/253,964
FILING DATE: 3 JUNE 1994
APPLICATION NUMBER: US 08/073,223
FILING DATE: 4 JUNE 1993
APPLICATION NUMBER: US 08/200,947
FILING DATE: 23 FEB 1994
APPLICATION NUMBER: US 07/864,805
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/247,505
FILING DATE: 23 MAY 1994
APPLICATION NUMBER: US 07/864,866
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 08/218,155
FILING DATE: 25 MAR 1994
APPLICATION NUMBER: US 07/864,807
FILING DATE: 7 APR 1992
APPLICATION NUMBER: US 07/902,467
FILING DATE: 16 JUNE 1992
APPLICATION NUMBER: US 07/275,433
FILING DATE: 23 NOV 1988
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-002CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1093
US-10-390-330-3

Query Match 100.0%; Score 20; DB 17; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 22
US-10-643-768-22/c
Sequence 22, Application US/10643768
Publication No. US20040192899A1
GENERAL INFORMATION:
APPLICANT: Sharpe, Arlene H.
APPLICANT: Borriello, Francescopaulo
APPLICANT: Freeman, Gordon J.
APPLICANT: Nadler, Lee M.

TITLE OF INVENTION: Novel Forms of T Cell Costimulatory
TITLE OF INVENTION: Molecules and Uses Therefor
FILE REFERENCE: BWI-120CPADV
CURRENT APPLICATION NUMBER: US/10/643,768
PRIOR FILING DATE: 2003-08-18
PRIOR APPLICATION NUMBER: US/09/837,867
PRIOR FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 08/205,697
PRIOR FILING DATE: 1994-03-02
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 1120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (107) ... (1093)
US-10-643-768-22

Query Match 100.0%; Score 20; DB 18; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 23

US-10-429-079B-1/c
Sequence 1, Application US/10429079B
Publication No. US20040230051A1

GENERAL INFORMATION:
APPLICANT: Freeman, Gordon J.
Nadler, Lee M.
Gray, Gary S.

TITLE OF INVENTION: THERAPEUTIC COMPOSITIONS FOR INHIBITING THE
INTERACTIONS OF B7-1 AND B7-2 WITH THEIR NATURAL
LIGANDS

NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street, 24th Floor
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/429,079B
FILING DATE: 02-May-2003
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/425,762
FILING DATE: 22-OCT-1999
APPLICATION NUMBER: 08/479,744
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/280,757
FILING DATE: 26-JUL-1994
APPLICATION NUMBER: 08/147,773
FILING DATE: 03-NOV-1993
APPLICATION NUMBER: 08/109,393
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: 08/101,624
FILING DATE: 26-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: RPI-004C3DVCN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1093
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-429-079B-1

Query Match 100.0%; Score 20; DB 18; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 24
US-10-444-206-296/c
; Sequence 296, Application US/10444206
; Publication No. US20040023917A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Clarence Frank
; APPLICANT: Vickers, Timothy A.
; APPLICANT: Karras, James G.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/444,206
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: 09/851,871
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/14471
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/326,186
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 08/777,266
; PRIOR FILING DATE: 1996-12-31
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 296
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (117)...(1106)
US-10-444-206-296

Query Match 100.0%; Score 20; DB 17; Length 2781;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 25
US-10-756-783-5/c
; Sequence 5, Application US/10756783
; Publication No. US20040209363A1
; GENERAL INFORMATION:
; APPLICANT: WATTS, Tania

APPLICANT: BUKCZYNSKI, Jacob
APPLICANT: WEN, Tao
; TITLE OF INVENTION: Methods and compositions for modulating T cell activation and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 12912/2
; CURRENT APPLICATION NUMBER: US/10/756,783
; PRIOR FILING DATE: 2004-01-12
; PRIOR APPLICATION NUMBER: US 60/304,430
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/CA02/01033
; PRIOR FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (117)...(1106)
US-10-756-783-5

Query Match 100.0%; Score 20; DB 18; Length 2781;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 26
US-10-641-962-296/c
; Sequence 296, Application US/10641962
; Publication No. US20040235164A1
; GENERAL INFORMATION:
; APPLICANT: Bennett et al.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE: 30566/39578
; CURRENT APPLICATION NUMBER: US/10/641,962
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 296
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (117)...(1106)
US-10-641-962-296

Query Match 100.0%; Score 20; DB 18; Length 2781;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
DB 67 CGTGTCTCTGTCTAGTCCC 48

RESULT 27
US-10-444-206-297/c
; Sequence 297, Application US/10444206
; Publication No. US20040023917A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Clarence Frank
; APPLICANT: Vickers, Timothy A.
; APPLICANT: Karras, James G.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; TITLE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE:

```
; CURRENT APPLICATION NUMBER: US/10/444,206
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: 09/851,871
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: PCT/US00/14471
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 09/326,186
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 08/777,266
; PRIOR FILING DATE: 1996 12 31
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 297
; LENGTH: 68001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-444-206-297

Query Match      100.0%; Score 20; DB 17; Length 68001;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 1067 CGTGTGCTGTGCTAGTCCC 1048

RESULT 28
US-10-641-962-297/c
; Sequence 297, Application US/10641962
; Publication No. US20040235164A1
; GENERAL INFORMATION:
; APPLICANT: Bennett et al.
; TITLE OF INVENTION: Oligonucleotide Compositions and Methods for the
; FILE OF INVENTION: Modulation of the Expression of B7 Protein
; FILE REFERENCE: 30566/39578
; CURRENT APPLICATION NUMBER: US/10/641,962
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 297
; LENGTH: 68001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-641-962-297

Query Match      100.0%; Score 20; DB 18; Length 68001;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 1067 CGTGTGCTGTGCTAGTCCC 1048

RESULT 29
US-09-796-692-7817/c
; Sequence 7817, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```

```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (535)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7817

Query Match      92.0%; Score 18.4; DB 9; Length 551;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 41 CGTGTGCTGTGCTAGTCCC 22

RESULT 30
US-10-040-862-7817/c
; Sequence 7817, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
```

```

; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (535)
; OTHER INFORMATION: n=A,T,C or G
; US-10-040-862-7817

```

```

Query Match          92.0%; Score 18.4; DB 14; Length 551;
Best Local Similarity 95.0%; Pred.No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CGTGTCTCTGTCTAGTCC 20
    ||||| ||||| |||||
DB 41 CGTGTGCGCTGTCTAGTCC 22

```

RESULT 31

```

US-10-057-475B-7817/c
; Sequence 7817, Application US/10057475B
; Publication No. US2004002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(551)
; OTHER INFORMATION: n = g, a, c or t
; US-10-154-884B-7817

```

Query Match

92.0%; Score 18.4; DB 17; Length 551;

```

; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(551)
; OTHER INFORMATION: n = g, a, c or t
; US-10-057-475B-7817

```

```

Query Match          92.0%; Score 18.4; DB 17; Length 551;
Best Local Similarity 95.0%; Pred.No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CGTGTCTCTGTCTAGTCC 20
    ||||| ||||| |||||
DB 41 CGTGTGCGCTGTCTAGTCC 22

```

RESULT 32

```

US-10-154-884B-7817/c
; Sequence 7817, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(551)
; OTHER INFORMATION: n = g, a, c or t
; US-10-154-884B-7817

```

Query Match

92.0%; Score 18.4; DB 17; Length 551;

```
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 41 CGTGTGCTGTGCTAGTCCC 22

RESULT 33
US-10-764-324-7817/c
; Sequence 7817, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7817
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (535)
; OTHER INFORMATION: n=A,T,C or G
US-10-764-324-7817

Query Match 92.0%; Score 18.4; DB 18; Length 551;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 41 CGTGTGCTGTGCTAGTCCC 22

RESULT 34
US-10-424-599-120635/c
; Sequence 120635, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1063
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 120635
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_79941C.1
US-10-424-599-120635

Query Match 87.0%; Score 17.4; DB 17; Length 215;
Best Local Similarity 94.7%; Pred. No. 65;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 144 GTGTGCTGTGCTAGTCCC 126

RESULT 35
US-10-767-701-18097
; Sequence 18097, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 18097
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-006-P1-K1-C2
US-10-767-701-18097

Query Match 84.0%; Score 16.8; DB 18; Length 319;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 20 CGTGTGCTGTGCAAGTGCC 39

RESULT 36
US-09-764-891-1063
; Sequence 1063, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1063
```

; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (298)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (368)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (372)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-1063

Query Match 80.0%; Score 16; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGTGTCGTGCTAGT 17
|||||
Db 70 GTGTGTCGTGCTAGT 85
|||||

RESULT 37

US-09-728-444-574
; Sequence 574, Application US/09728444
; Patent No. US20020161207A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020161207A1 Murine Polynucleotide Sequences
; FILE REFERENCE: LEX-0100-USA
; CURRENT APPLICATION NUMBER: US/09/728,444
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,360
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1206
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(515)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-444-574

Query Match 80.0%; Score 16; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.3e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGT 16
|||||
Db 384 CGTGTGTCGTGCTAGT 399
|||||

RESULT 38

US-10-087-192-613/c
; Sequence 613, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 613
; LENGTH: 41434
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(41434)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-613

Query Match 80.0%; Score 16; DB 13; Length 41434;
Best Local Similarity 100.0%; Pred. No. 3.4e+02; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGT 16
|||||
Db 5400 CGTGTGTCGTGCTAGT 5385
|||||

RESULT 39

US-10-719-900-719033
; Sequence 719033, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 719033
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-719033

Query Match 79.0%; Score 15.8; DB 19; Length 25;
Best Local Similarity 89.5%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGTCGTGCTAGTCCC 20
|||||
Db 1 GTGTGTCGTGCTAGACCC 19
|||||

RESULT 40

US-10-674-124A-23858/c
; Sequence 23858, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: GENETIC POLYMORPHISM MARKERS
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28

```
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 23858
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC008576.6_117929
; FEATURE:
; OTHER INFORMATION: Located on chromosome 19
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 10025297
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 513
US-10-674-124A-23858

Query Match          79.0%; Score 15.8; DB 18; Length 293;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GTGTGCTCTGTGCTAGTCCC 20
Db      187 GTGTGCTCTGTGCTATATCCC 169

RESULT 41
US-10-276-774-1153
; Sequence 1153, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1153
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = a,t,c or g
US-10-276-774-1153

Query Match          79.0%; Score 15.8; DB 17; Length 439;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGTGTGCTCTGTGCTAGTCC 19
Db      39 CGTGTGCTCTGTGCGAGGCC 57

RESULT 42
US-10-632-290491
; Sequence 290491, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290491
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-290491

Query Match          79.0%; Score 15.8; DB 13; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GTGTGCTCTGTGCTAGTCCC 20
Db      403 GTGTTCCTGTGCCAGTCCC 421

RESULT 43
US-10-027-632-290492
; Sequence 290492, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290492
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-290492

Query Match          79.0%; Score 15.8; DB 13; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GTGTGCTCTGTGCTAGTCCC 20
```



```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146229
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146229

Query Match      79.0%; Score 15.8; DB 17; Length 757;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTGCTAGTCC 19
DB 567 CGTGTGCTCTGTGCGAGGCC 585

RESULT 48
US-10-027-632-164600/c
; Sequence 164600, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164600
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164600

Query Match      79.0%; Score 15.8; DB 13; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTGCTAGTCC 19
DB 529 CGTGTGCTCTGTGCGAGGCC 511

RESULT 49
US-10-027-632-164601/c
; Sequence 164601, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164601
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164601

Query Match      79.0%; Score 15.8; DB 13; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTCTGTGCTAGTCC 19
DB 529 CGTGTGCTCTGTGCGAGGCC 511

RESULT 50
US-10-027-632-164600/c
; Sequence 164600, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164600
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164600

Query Match      79.0%; Score 15.8; DB 17; Length 804;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 CGTGTCTCTGTCTAGTCC 19
|||
Db 529 CGTGTCTCTGTCTAGTCC 511
|||

Search completed: March 15, 2005, 22:52:08
Job time : 518 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 18:05:05 ; Search time 127 Seconds
(without alignments)
257.681 Million cell updates/sec

Title: US-09-980-953-256
Perfect score: 20
Sequence: 1 cgtgtgtctgtctgtatgcc 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Issued Patents NA.*
1: /cgn2.6/prodata/1/ina/5A COMB.seq.*
2: /cgn2.6/prodata/1/ina/5B COMB.seq.*
3: /cgn2.6/prodata/1/ina/6A COMB.seq.*
4: /cgn2.6/prodata/1/ina/6B COMB.seq.*
5: /cgn2.6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2.6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	124	3	US-08-205-6978-39
C 2	20	100.0	124	3	US-08-702-525-39
C 3	20	100.0	124	4	US-09-837-8678-39
C 4	20	100.0	124	5	PCT-US95-02576-39
C 5	20	100.0	430	4	US-09-513-999C-425
C 6	20	100.0	1112	4	US-09-441-411-25
C 7	20	100.0	1120	2	US-08-456-104-1
C 8	20	100.0	1120	2	US-08-101-624-1
C 9	20	100.0	1120	3	US-08-479-744A-1
C 10	20	100.0	1120	3	US-08-280-757B-1
C 11	20	100.0	1120	3	US-08-205-6978-22
C 12	20	100.0	1120	3	US-08-702-525-22
C 13	20	100.0	1120	3	US-08-403-253A-3
C 14	20	100.0	1120	4	US-08-435-816A-3
C 15	20	100.0	1120	4	US-09-425-762-1
C 16	20	100.0	1120	4	US-09-837-8678-22
C 17	20	100.0	1120	4	US-09-206-132-1
C 18	20	100.0	1120	4	US-09-425-516-1
C 19	20	100.0	1120	5	PCT-US95-02576-22
C 20	16.4	82.0	264206	4	US-09-949-016-12731
C 21	16.4	82.0	264304	4	US-09-949-016-13249
C 22	16	80.0	368	4	US-09-513-999C-9101
C 23	15.8	79.0	2304	4	US-09-799-451-340
C 24	15.8	79.0	26007	4	US-09-949-016-17222
C 25	15.8	79.0	85850	4	US-09-949-016-13424
C 26	15.8	79.0	151295	4	US-09-949-016-14568
C 27	15.8	79.0	151295	4	US-09-949-016-14569

c 393 13.8 69.0 58768 4 US-09-949-016-13175 Sequence 13175, A
394 13.8 69.0 59719 4 US-09-949-016-15616 Sequence 15616, A
395 13.8 69.0 61178 4 US-09-949-016-17369 Sequence 17369, A
396 13.8 69.0 61664 4 US-09-949-016-13308 Sequence 13308, A
397 13.8 69.0 70308 4 US-09-949-016-15501 Sequence 15501, A
398 13.8 69.0 75929 4 US-09-949-016-15643 Sequence 15643, A
399 13.8 69.0 75929 4 US-09-949-016-15608 Sequence 15608, A
400 13.8 69.0 77626 4 US-09-949-016-12548 Sequence 12548, A
401 13.8 69.0 83617 4 US-09-949-016-12254 Sequence 12254, A
402 13.8 69.0 84839 4 US-09-949-016-15816 Sequence 15816, A
403 13.8 69.0 87190 4 US-09-949-016-16335 Sequence 16335, A
c 404 13.8 69.0 89210 4 US-09-949-016-16988 Sequence 16988, A
c 405 13.8 69.0 91665 4 US-09-949-016-12234 Sequence 12234, A
c 406 13.8 69.0 95621 4 US-09-949-016-13237 Sequence 13237, A
407 13.8 69.0 97195 4 US-09-949-016-12212 Sequence 12212, A
408 13.8 69.0 97196 4 US-09-949-016-16971 Sequence 16971, A
409 13.8 69.0 100928 4 US-09-949-016-16926 Sequence 16926, A
410 13.8 69.0 113100 4 US-09-949-016-12245 Sequence 12245, A
411 13.8 69.0 113876 4 US-09-949-016-14828 Sequence 14828, A
412 13.8 69.0 113876 4 US-09-949-016-14829 Sequence 14829, A
413 13.8 69.0 115508 4 US-09-949-016-11800 Sequence 11800, A
414 13.8 69.0 115508 4 US-09-949-016-14826 Sequence 14826, A
c 415 13.8 69.0 115508 4 US-09-949-016-14827 Sequence 14827, A
c 416 13.8 69.0 133063 4 US-09-949-016-16293 Sequence 16293, A
c 417 13.8 69.0 155266 4 US-09-949-016-13870 Sequence 13870, A
418 13.8 69.0 161124 4 US-09-949-016-11760 Sequence 11760, A
c 419 13.8 69.0 175236 4 US-09-949-016-14353 Sequence 14353, A
420 13.8 69.0 211049 4 US-09-949-016-15770 Sequence 15770, A
c 421 13.8 69.0 211049 4 US-09-949-016-15770 Sequence 15770, A
422 13.8 69.0 236474 4 US-09-949-016-13418 Sequence 13418, A
423 13.8 69.0 283538 4 US-09-949-016-13506 Sequence 13506, A
424 13.8 69.0 298336 4 US-09-949-016-16600 Sequence 16600, A
425 13.8 69.0 312957 4 US-09-949-001-31 Sequence 31, Appl
426 13.8 69.0 312972 4 US-09-949-001-34 Sequence 34, Appl
427 13.8 69.0 343352 4 US-09-949-016-13498 Sequence 13498, A
428 13.8 69.0 392000 4 US-10-027-983-11 Sequence 11, Appl
429 13.8 69.0 450395 4 US-09-949-016-15473 Sequence 15473, A
c 430 13.6 68.0 106 3 US-08-482-677-13 Sequence 13, Appl
c 431 13.6 68.0 106 4 US-10-033-174-13 Sequence 13, Appl
432 13.6 68.0 187 4 US-09-270-767-3794 Sequence 3794, Ap
c 433 13.6 68.0 187 4 US-09-270-767-19076 Sequence 19076, A
c 434 13.6 68.0 231 4 US-09-902-540-1610 Sequence 1610, Ap
c 435 13.6 68.0 310 4 US-09-513-999C-30296 Sequence 30296, A
c 436 13.6 68.0 329 4 US-09-621-976-7848 Sequence 7848, Ap
c 437 13.6 68.0 331 4 US-09-513-999C-32221 Sequence 32221, A
438 13.6 68.0 367 3 US-08-908-223-180 Sequence 180, App
c 439 13.6 68.0 415 4 US-09-270-767-29304 Sequence 29304, A
c 440 13.6 68.0 416 4 US-09-621-976-1336 Sequence 136, App
c 441 13.6 68.0 443 4 US-09-621-976-14153 Sequence 14153, A
c 442 13.6 68.0 443 4 US-09-270-767-3949 Sequence 3949, Ap
c 443 13.6 68.0 443 4 US-09-270-767-19231 Sequence 19231, A
c 444 13.6 68.0 566 4 US-09-949-016-83728 Sequence 83728, A
c 445 13.6 68.0 601 4 US-09-949-016-22192 Sequence 22192, A
c 446 13.6 68.0 601 4 US-09-949-016-22193 Sequence 22193, A
c 447 13.6 68.0 601 4 US-09-949-016-26341 Sequence 26341, A
c 448 13.6 68.0 601 4 US-09-949-016-26342 Sequence 26342, A
c 449 13.6 68.0 601 4 US-09-949-016-68800 Sequence 68800, A
c 450 13.6 68.0 601 4 US-09-949-016-78595 Sequence 78595, A
c 451 13.6 68.0 601 4 US-09-949-016-83729 Sequence 83729, A
c 452 13.6 68.0 601 4 US-09-949-016-83730 Sequence 83730, A
c 453 13.6 68.0 601 4 US-09-949-016-91772 Sequence 91772, A
c 454 13.6 68.0 601 4 US-09-949-016-105720 Sequence 105720, A
c 455 13.6 68.0 601 4 US-09-949-016-105728 Sequence 105728, A
c 456 13.6 68.0 601 4 US-09-949-016-130368 Sequence 130368, A
c 457 13.6 68.0 601 4 US-09-949-016-140995 Sequence 140995, A
c 458 13.6 68.0 601 4 US-09-949-016-140996 Sequence 140996, A
c 459 13.6 68.0 601 4 US-09-949-016-156013 Sequence 156013, A
c 460 13.6 68.0 601 4 US-09-949-016-170283 Sequence 170283, A
c 461 13.6 68.0 601 4 US-09-949-016-176149 Sequence 176149, A
c 462 13.6 68.0 601 4 US-09-949-016-189048 Sequence 189048, A
c 463 13.6 68.0 601 4 US-09-949-016-189049 Sequence 189049, A
c 464 13.6 68.0 601 4 US-09-949-016-194713 Sequence 194713, A
c 465 13.6 68.0 601 4 US-09-949-016-203029 Sequence 203029, A

Sequence 203037, App
Sequence 205849, App
Sequence 205850, App
Sequence 109, App
Sequence 628, App
Sequence 232, App
Sequence 157, App
Sequence 158, App
Sequence 13352, A
Sequence 4554, Ap
Sequence 5, Appli
Sequence 236, App
Sequence 138, App
Sequence 3, Appli
Sequence 1, Appli
Sequence 7089, Ap
Sequence 9, Appli
Sequence 760, App
Sequence 1, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 15, Appl
Sequence 13, Appl
Sequence 253, App
Sequence 5522, App
Sequence 713, App
Sequence 38, Appl
Sequence 38, Appl
Sequence 246, App
Sequence 11, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 5809, Ap
Sequence 207, App

13.6 68.0 601 4 US-09-949-016-203037
13.6 68.0 601 4 US-09-949-016-205849
13.6 68.0 601 4 US-09-949-016-205850
13.6 68.0 716 4 US-09-023-655-109
13.6 68.0 720 3 US-08-998-416-628
13.6 68.0 905 4 US-09-976-594-232
13.6 68.0 1001 4 US-09-671-317-157
13.6 68.0 1001 4 US-09-671-317-158
13.6 68.0 1007 4 US-09-270-767-13352
13.6 68.0 1057 4 US-09-949-016-4554
13.6 68.0 1119 4 US-09-784-358-5
13.6 68.0 1244 4 US-09-902-540-236
13.6 68.0 1289 3 US-09-247-155-138
13.6 68.0 1341 4 US-09-784-358-3
13.6 68.0 1430 3 US-09-492-985-1
13.6 68.0 1797 4 US-09-489-039A-7089
13.6 68.0 1953 4 US-09-784-358-9
13.6 68.0 1969 4 US-09-799-451-760
13.6 68.0 2005 1 US-08-208-108-1
13.6 68.0 2175 4 US-09-784-358-7
13.6 68.0 2181 3 US-09-126-980-1
13.6 68.0 2181 3 US-09-476-482-1
13.6 68.0 2181 3 US-09-517-605-15
13.6 68.0 2316 4 US-09-784-358-13
13.6 68.0 2332 4 US-09-949-016-253
13.6 68.0 2332 4 US-09-949-016-5522
13.6 68.0 2425 4 US-09-799-451-713
13.6 68.0 2517 4 US-09-894-988A-38
13.6 68.0 2517 4 US-10-237-551-38
13.6 68.0 2517 4 US-10-237-551-246
13.6 68.0 2538 4 US-09-784-358-11
13.6 68.0 2931 4 US-09-623-624-1
13.6 68.0 2931 4 US-10-270-595-1
13.6 68.0 2934 4 US-09-949-016-5809
13.6 68.0 2943 4 US-09-949-016-207

ALIGNMENTS

RESULT 1
US-08-205-697A-39/c
; Sequence 39, Application US/08205697A
; Patent No. 6218510
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules
; TITLE OF INVENTION: and Uses Therefor
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,697A
FILING DATE: 02-Mar-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400

```

; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-205-697A-39

Query Match      100.0%; Score 20; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 3
US-09-837-867A-39/c
; Sequence 39, Application US/09837867A
; Patent No. 6608180
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 6608180el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BWI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(124)
US-09-837-867A-39

Query Match      100.0%; Score 20; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 4
PCT-US95-02576-39/c
; Sequence 39, Application PC/TUS9502576
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
; TITLE OF INVENTION: and Uses Therefor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION NUMBER: US/08/205,697
; FILING DATE: 02-Mar-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-702-525-39

Query Match      100.0%; Score 20; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGTGTCTCTGTCTAGTCCC 20
DB      57 CGTGTCTCTGTCTAGTCCC 38

RESULT 2
US-08-702-525-39/c
; Sequence 39, Application US/08702525
; Patent No. 6294660
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaolo
; APPLICANT: Freeman, Gordon
; APPLICANT: Nadler, Lee
; TITLE OF INVENTION: No. 6294660el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION NUMBER: US/08/702,525
; FILING DATE:
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/205,697
; FILING DATE: 02-Mar-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: BWI-120CPUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
US-08-702-525-39
```

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..124
PCT-US95-02576-39

Query Match 100.0%; Score 20; DB 5; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 5
US-09-513-999C-425/c
; Sequence 425, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59-US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 425
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 125..430
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 125..199
; OTHER INFORMATION: score 7.1
; OTHER INFORMATION: seq LFVMAFLLSGAAP/LK
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 116
; OTHER INFORMATION: r-a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 121
; OTHER INFORMATION: m-a or c

Query Match 100.0%; Score 20; DB 4; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 75 CGTGTGCTGTGCTAGTCCC 56

, RESULT 6

US-09-441-411-25/c
; Sequence 25, Application US/09441411
; Patent No. 6734172
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-25

Query Match 100.0%; Score 20; DB 4; Length 1112;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 7
US-08-456-104-1/c
; Sequence 1, Application US/08456104
; Patent No. 5861310
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: TUMOR CELLS MODIFIED TO EXPRESS B7-2 AND B7-3 WITH INCREASED
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/456,104
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/101,624;
; FILING DATE: 26-JUL-1993;
; APPLICATION NUMBER: 08/109,393;
; APPLICATION NUMBER: 19-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-456-104-1
Query Match 100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
Db 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 8
US-08-101-624-1/c
; Sequence 1, Application US/08101624
; Patent No. 5942607
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: No. 5942607el CTLA4/CD28 Ligands and
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/101,624
; FILING DATE: 26-JUL-1993
; PRIOR APPLICATION NUMBER:
; CLASSIFICATION: 514
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-101-624-1
Query Match 100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
Db 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 9
US-08-479-744A-1/c
; Sequence 1, Application US/08479744A
; Patent No. 6084067
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: No. 6084067el CTLA4/CD28 Ligands and
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,744A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/280,757
; FILING DATE: 26-JUL-1994
; APPLICATION NUMBER: 08/109,393
; FILING DATE: 28-AUG-1993
; APPLICATION NUMBER: 08/101,624
; FILING DATE: 26-JULY-1993
; APPLICATION NUMBER: 08/147,773
; FILING DATE: 3-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-479-744A-1
Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
Db 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 10
US-08-280-757B-1/c
; Sequence 1, Application US/08280757B
; Patent No. 6130316
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; APPLICANT: Greenfield, Edward
```

;/ TITLE OF INVENTION: No. 6130316el CTLA4/CD28 Ligands and
;/ TITLE OF INVENTION: Uses Therefor
;/ NUMBER OF SEQUENCES: 53
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: LAHIVE & COCKFIELD
;/ STREET: 60 State Street, Suite 510
;/ CITY: Boston
;/ STATE: Massachusetts
;/ COUNTRY: USA
;/ ZIP: 02109

;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/280,757B
;/ FILING DATE: 26-JUL-1994
;/ CLASSIFICATION: 435

;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 08/101,624
;/ FILING DATE: 26-JULY-1993
;/ APPLICATION NUMBER: 08/109,393
;/ FILING DATE: 19-AUG-1993
;/ APPLICATION NUMBER: 08/147,773
;/ FILING DATE: 3-NOV-1993

;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Mandragouras, Amy E.
;/ REGISTRATION NUMBER: 36,207
;/ REFERENCE/DOCKET NUMBER: RPI-004CP2
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (617) 227-7400
;/ TELEFAX: (617) 227-5941

;/ INFORMATION FOR SEQ ID NO: 1:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1120 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: CDNA
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: 107..1093

US-08-280-757B-1

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11
US-08-205-697A-22/c
; Sequence 22, Application US/08205697A
; Patent No. 6218510
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.

;/ TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules
;/ NUMBER OF SEQUENCES: 61
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: LAHIVE & COCKFIELD
;/ STREET: 60 State Street, suite 510
;/ CITY: Boston
;/ STATE: Massachusetts
;/ COUNTRY: USA
;/ ZIP: 02109-1875

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11
US-08-205-697A-22/c
; Sequence 22, Application US/08205697A
; Patent No. 6218510
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.

;/ TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules
;/ NUMBER OF SEQUENCES: 61
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: LAHIVE & COCKFIELD
;/ STREET: 60 State Street, suite 510
;/ CITY: Boston
;/ STATE: Massachusetts
;/ COUNTRY: USA
;/ ZIP: 02109-1875

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11
US-08-205-697A-22/c
; Sequence 22, Application US/08205697A
; Patent No. 6218510
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.

;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: ASCII Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/205,697A
;/ FILING DATE: 02-Mar-1994
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Mandragouras, Amy E.
;/ REGISTRATION NUMBER: 36,207
;/ REFERENCE/DOCKET NUMBER: BWI-120
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (617)227-7400
;/ TELEFAX: (617)227-5941

;/ INFORMATION FOR SEQ ID NO: 22:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 1120 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: CDNA
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: 107..1093

US-08-205-697A-22

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 12
US-08-702-525-22/c
; Sequence 22, Application US/08702525
; Patent No. 6294860
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Sharpe
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon
; APPLICANT: Nadler, Lee
; TITLE OF INVENTION: No. 6294660el Forms of T Cell Costimulatory
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: ASCII Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/702,525
;/ FILING DATE:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/205,697
;/ FILING DATE: 02-Mar-1994
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Mandragouras, Amy E.
;/ REGISTRATION NUMBER: 36,207
;/ REFERENCE/DOCKET NUMBER: BWI-120CPUS
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (617)227-7400
;/ TELEFAX: (617)227-5941

; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-702-525-22

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTAGTCC 20
Db 57 CGTGTCTGTCTAGTCC 38

RESULT 13

US-08-403-253A-3/c
; Sequence 3, Application US/08403253A
; Patent No. 6352694
; GENERAL INFORMATION:
; APPLICANT: June, Carl H., Thompson, Craig B., Nabel, Gary J.
; APPLICANT: Gray, Gary S., Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,253A
; FILING DATE: March 10, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,964
; FILING DATE: 3 JUNE 1994
; APPLICATION NUMBER: US 08/073,223
; FILING DATE: 4 JUNE 1993
; APPLICATION NUMBER: US 08/200,947
; FILING DATE: 23 FEB 1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/247,505
; FILING DATE: 23 MAY 1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25 MAR 1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16 JUNE 1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23 NOV 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
US-08-403-253A-3

Query Match 100.0%; Score 20; DB 3; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTGTCTAGTCC 20
Db 57 CGTGTCTGTCTAGTCC 38

RESULT 14

US-08-435-816A-3/c
; Sequence 3, Application US/08435816A
; Patent No. 6534055
; GENERAL INFORMATION:
; APPLICANT: June, Carl H.
; APPLICANT: Thompson, Craig B.
; APPLICANT: Nabel, Gary J.
; APPLICANT: Gray, Gary S.
; APPLICANT: Rennert, Paul D.
; TITLE OF INVENTION: Methods For Selectively Stimulating Proliferation Of T-Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,816A
; FILING DATE: May 4, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,253
; FILING DATE: 10 MARCH 1995
; APPLICATION NUMBER: US 08/253,964
; FILING DATE: 3 JUNE 1994
; APPLICATION NUMBER: US 08/073,223
; FILING DATE: 4 JUNE 1993
; APPLICATION NUMBER: US 08/200,947
; FILING DATE: 23 FEB 1994
; APPLICATION NUMBER: US 07/864,805
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/247,505
; FILING DATE: 23 MAY 1994
; APPLICATION NUMBER: US 07/864,866
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 08/218,155
; FILING DATE: 25 MAR 1994
; APPLICATION NUMBER: US 07/864,807
; FILING DATE: 7 APR 1992
; APPLICATION NUMBER: US 07/902,467
; FILING DATE: 16 JUNE 1992
; APPLICATION NUMBER: US 07/275,433
; FILING DATE: 23 NOV 1988

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-002CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; US-08-435-816A-3

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATGCC 20
Db 57 CGTGTGCTGCTAGTATGCC 38

RESULT 15
US-09-425-762-1/c
; Sequence 1, Application US/09425762
; Patent No. 6605279
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: No. 6605279el CTLA4/CD28 Ligands and
; TITLE OF INVENTION: Uses Therefor
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,762
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-004CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..1093
; US-09-425-762-1

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATGCC 20
Db 57 CGTGTGCTGCTAGTATGCC 38

RESULT 16
US-09-837-867A-22/c
; Sequence 22, Application US/09837867A
; Patent No. 6608180
; GENERAL INFORMATION:
; APPLICANT: Sharpe, Arlene H.
; APPLICANT: Borriello, Francescopaulo
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; TITLE OF INVENTION: No. 8608180el Forms of T Cell Costimulatory
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: BMI-120CPADV
; CURRENT APPLICATION NUMBER: US/09/837,867A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 08/205,697
; PRIOR FILING DATE: 1994-03-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(1093)
; US-09-837-867A-22

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATGCC 20
Db 57 CGTGTGCTGCTAGTATGCC 38

RESULT 17
US-09-206-132-1/c
; Sequence 1, Application US/09206132
; Patent No. 6723705
; GENERAL INFORMATION:
; APPLICANT: Freeman, Gordon J.
; APPLICANT: Nadler, Lee M.
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: TUMOR CELLS MODIFIED TO EXPRESS B7-2 AND B7-3
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/206,132
FILING DATE: 08/147,773
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/456,104
FILING DATE: 08/101,624
APPLICATION NUMBER: 08/101,624
FILING DATE: 26-JUL-1993
APPLICATION NUMBER: 08/109,393
APPLICATION NUMBER: 19-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1093
US-09-206-132-1

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||||
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 18
PCT-US95-425-516-1/c
Sequence 1, Application US/09425516
Patent No. 6824779
GENERAL INFORMATION:
APPLICANT: Freeman, Gordon J.
APPLICANT: Nadler, Lee M.
APPLICANT: Gray, Gary S.
TITLE OF INVENTION: No. 6824779el CTLA4/CD28 Ligands and
TITLE OF INVENTION: Uses therefor
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/425,516
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,744
FILING DATE: June 7, 1995
APPLICATION NUMBER: 08/280,757
FILING DATE: 26-JUL-1994
APPLICATION NUMBER: 08/109,393
FILING DATE: 28-AUG-1993
APPLICATION NUMBER: 08/101,624

FILING DATE: 26-JULY-1993
APPLICATION NUMBER: 08/147,773
FILING DATE: 3-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: RPI-004CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1093
US-09-425-516-1

Query Match 100.0%; Score 20; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTCTAGTCCC 20
|||||
DB 57 CGTGTCTCTGTCTAGTCCC 38

RESULT 19
PCT-US95-02576-22/c
Sequence 22, Application PC/TUS9502576
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
TITLE OF INVENTION: and Uses Therefor
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02576
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,697
FILING DATE: 02-Mar-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120CPPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS

```
; LOCATION: 107..1093
PCT-US95-02576-22

Query Match      100.0%; Score 20; DB 5; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGTGTCTGTGCTAGTCC 20
Db 57 CCGTGTCTGTGCTAGTCC 38

RESULT 20
US-09-949-016-12731
; Sequence 12731, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12731
; LENGTH: 264206
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12731

Query Match      82.0%; Score 16.4; DB 4; Length 264206;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGTCTGTGCTAGTCC 19
Db 37623 GTGTGTCTGTGCTAGTCC 37640

RESULT 21
US-09-949-016-13249
; Sequence 13249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 264304
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13249

Query Match      82.0%; Score 16.4; DB 4; Length 264304;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGTCTGTGCTAGTCC 19
Db 37623 GTGTGTCTGTGCTAGTCC 37640

RESULT 22
US-09-513-999C-9101
; Sequence 9101, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 9101
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 202
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 268
; OTHER INFORMATION: k=g or t
; NAME/KEY: misc_feature
; LOCATION: 323
; OTHER INFORMATION: d=a or g or t
US-09-513-999C-9101

Query Match      80.0%; Score 16; DB 4; Length 368;
Best Local Similarity 88.9%; Pred. No. 1.3e+02;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGTCTGTGCTAGTCC 19
Db 192 GTGTGTCTGTGCTAGTCC 209

RESULT 23
US-09-799-451-340
; Sequence 340, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
```

; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 340
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1974)
US-09-799-451-340

Query Match 79.0%; Score 15.8; DB 4; Length 2304;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGTGCTAGTCC 19
|||||
DB 651 CGTGTCTCTGTGCGAGGCC 669

RESULT 24
US-09-949-016-17222
; Sequence 17222, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17222
; LENGTH: 26007
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17222

Query Match 79.0%; Score 15.8; DB 4; Length 26007;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGTGTCTGTGCTAGTCC 20
|||||
DB 25893 GGGTGTCTGGGCTAGTCC 25911

RESULT 25
US-09-949-016-13424/c
; Sequence 13424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13424
; LENGTH: 85850
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(85850)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13424

Query Match 79.0%; Score 15.8; DB 4; Length 85850;
Best Local Similarity 89.5%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20
|||||
DB 6643 GTGTGTGTGCTATTCCC 6625

RESULT 26
US-09-949-016-14568
; Sequence 14568, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14568
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14568

Query Match 79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20
|||||
DB 143140 GTGTGTCTGTGCCAGTGCC 143158

RESULT 27
US-09-949-016-14569
; Sequence 14569, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14569
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14569

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

RESULT 28
US-09-949-016-14570
; Sequence 14570, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14570
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14570

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

RESULT 29
US-09-949-016-14571
; Sequence 14571, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14571
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14571

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

RESULT 30
US-09-949-016-14572
; Sequence 14572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14572
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14572

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

RESULT 31
US-09-949-016-14573
; Sequence 14573, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14573
; LENGTH: 151295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14573

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14571
LENGTH: 151295
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(151295)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14571

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14570
LENGTH: 151295
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(151295)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14570

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14572
LENGTH: 151295
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(151295)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14572

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14573
LENGTH: 151295
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(151295)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14573

Query Match          79.0%; Score 15.8; DB 4; Length 151295;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
Db 143140 GTGTGCTGTGCCAGTGCC 143158
```

US-09-949-016-14573

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14573
; LENGTH: 393753
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(393753)
; OTHER INFORMATION: n = A,T,C or G

Query Match 79.0%; Score 15.8; DB 4; Length 393753;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
|||||

Db 385598 GTGTCTCTGTCTAGTCCC 385616

RESULT 32

US-09-949-016-14574

Sequence 14574, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14574
; LENGTH: 393753
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(393753)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14574

Query Match 79.0%; Score 15.8; DB 4; Length 393753;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
|||||

Db 385598 GTGTCTCTGTCTAGTCCC 385616

US-09-949-016-14546

Sequence 14546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14546
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14546

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
|||||

Db 809973 GTGTCTCTGTCTAGTCCC 809991

RESULT 34

US-09-949-016-14547

Sequence 14547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14547
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14547

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
|||||

Db 809973 GTGTCTCTGTCTAGTCCC 809991

```
RESULT 35
US-09-949-016-14548
; Sequence 14548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14548
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14548

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCC 20
Db      809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 36
US-09-949-016-14549
; Sequence 14549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14549
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCC 20
Db      809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||
```

```
Qy      2 GTGTGCTGTGCTAGTCC 20
Db      809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 37
US-09-949-016-14550
; Sequence 14550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14550
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14550

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GTGTGCTGTGCTAGTCC 20
Db      809973 GTGTGCTGTGCCAGTGCC 809991
|||||
|||||

RESULT 38
US-09-949-016-14551
; Sequence 14551, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14551
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14551

Query Match          79.0%; Score 15.8; DB 4; Length 818128;
```

```
Best Local Similarity 89.5%; Pred. No. 3.4e+02; DB 4; Length 818128;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 39
US-09-949-016-14552
; Sequence 14552, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14552
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14552

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 40
US-09-949-016-14553
; Sequence 14553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14553
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14553
```

```
US-09-949-016-14553
Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 41
US-09-949-016-14554
; Sequence 14554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14554

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGTAGTCC 20
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 42
US-09-949-016-14555
; Sequence 14555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14555
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14555
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14555

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 43
US-09-949-016-14556
; Sequence 14556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14556
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14556

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 44
US-09-949-016-14557
; Sequence 14557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14557
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14557

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 45
US-09-949-016-14558
; Sequence 14558, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14558
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14558

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991

RESULT 46
US-09-949-016-14559
; Sequence 14559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14559
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14559

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCC 20
   |||||
Db 809973 GTGTGCTGTGCCAGTGCC 809991
```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14559
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14559

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 47
US-09-949-016-14560
; Sequence 14560, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14560
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14560

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 48
US-09-949-016-14561
; Sequence 14561, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14561
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14561

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 49
US-09-949-016-14562
; Sequence 14562, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14562
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14562

Query Match 79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20
Db 809973 GTGTCTCTGTCCAGTGCC 809991

RESULT 50
US-09-949-016-14564
; Sequence 14564, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14564
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14564

Query Match      79.0%; Score 15.8; DB 4; Length 818128;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  GGTGTCTGTGCTAGTCCC 20
          |||||
Db      809973  GTGTCTGTGCTAGTCCC 809991
```

Search completed: March 15, 2005, 21:31:30
Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 17:31:56 ; Search time 3113 Seconds
(without alignments)
244.550 Million cell updates/sec

Title: US-09-980-953/256

Perfect score: 20
Sequence: 1 cgtgtgtctgtgtagtccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	220	6	CD523014 AGENCOURT
C 2	20	100.0	490	6	CD698408 EST14931
C 3	20	100.0	626	5	BX371662 BX371662
C 4	20	100.0	690	5	BQ109523 imageqc 7
C 5	20	100.0	695	4	BI767024 603054234
C 6	20	100.0	709	5	BQ109553 imageqc 7
C 7	19	95.0	132	6	CD691973 EST8496 h
C 8	19	95.0	753	4	BI906246 603063172
C 9	18.4	92.0	655	4	BI824940 603032554
C 10	18.4	92.0	995	7	CO847434 ILLUMINEX
C 11	17.4	87.0	542	1	AI992726 701493407
C 12	17.4	87.0	725	9	CC488528 CH240_320
C 13	17.4	87.0	864	8	BZ166194 CH230-463
C 14	17.4	87.0	1126	2	BF304344 CH1887245
C 15	17.4	87.0	2004	9	CL508080 SAIL 792
C 16	17	85.0	1245	5	BUI95329 AGENCOURT
C 17	16.8	84.0	363	6	BY670167 BY670167
C 18	16.8	84.0	395	9	BX963297 Reverse s
C 19	16.8	84.0	571	7	CF805725 psHA0071A
C 20	16.8	84.0	615	9	BX968677 Forward s
C 21	16.8	84.0	659	8	BZ285459 CH230-336
C 22	16.8	84.0	666	2	AW339267 xz89f02.x
C 23	16.8	84.0	696	9	CR120951 Forward s
C 24	16.8	84.0	760	9	AG353788 Mus muscu

25	16.8	84.0	840	4	BI833379
C 26	16.8	84.0	880	7	CK017551
C 27	16.8	84.0	882	6	CD754743
C 28	16.8	84.0	928	9	CR811620
C 29	16.8	84.0	937	6	CD252254
C 30	16.8	84.0	972	9	AG426940
C 31	16.8	84.0	994	4	BM048685
C 32	16.8	84.0	1010	6	CB994004
C 33	16.8	84.0	1028	5	BQ707308
C 34	16.8	84.0	1263	4	BI414634
C 35	16.8	84.0	1560	7	CF106699
C 36	16.4	82.0	178	1	AU076753
C 37	16.4	82.0	255	4	BM011282
C 38	16.4	82.0	411	7	CR546210
C 39	16.4	82.0	451	4	BG408744
C 40	16.4	82.0	481	2	AW281714
C 41	16.4	82.0	497	8	BH111553
C 42	16.4	82.0	500	6	CD589669
C 43	16.4	82.0	517	1	AA067552
C 44	16.4	82.0	544	6	CA041772
C 45	16.4	82.0	578	4	BM089797
C 46	16.4	82.0	592	6	CB504055
C 47	16.4	82.0	630	6	CA041823
C 48	16.4	82.0	636	6	CA061753
C 49	16.4	82.0	648	6	CB505188
C 50	16.4	82.0	657	1	AL847928
C 51	16.4	82.0	663	6	CA379658
C 52	16.4	82.0	684	4	BJ594673
C 53	16.4	82.0	711	4	BJ590911
C 54	16.4	82.0	730	1	AJ425708
C 55	16.4	82.0	732	4	BJ587908
C 56	16.4	82.0	756	4	BJ579066
C 57	16.4	82.0	756	4	BJ610428
C 58	16.4	82.0	784	5	BP142881
C 59	16.4	82.0	787	6	CA362363
C 60	16.4	82.0	787	9	AG402582
C 61	16.4	82.0	791	9	CNS0376M
C 62	16.4	82.0	798	9	CR080093
C 63	16.4	82.0	808	5	BP170401
C 64	16.4	82.0	817	9	CNS038AK
C 65	16.4	82.0	833	7	CNI163617
C 66	16.4	82.0	843	9	CNS01VIB
C 67	16.4	82.0	870	9	CNS031S7
C 68	16.4	82.0	1007	2	BF537884
C 69	16.4	82.0	1185	4	BG293564
C 70	16.4	82.0	1223	9	AG574348
C 71	16.4	82.0	1366	9	AG321235
C 72	16.4	82.0	1686	4	BG4848292
C 73	16.4	82.0	1691	2	BF342971
C 74	16	80.0	267	8	AZ296490
C 75	16	80.0	422	9	CR224707
C 76	16	80.0	445	8	AZ625316
C 77	16	80.0	450	8	AG680838
C 78	16	80.0	612	8	AZ427072
C 79	16	80.0	829	2	BE569899
C 80	16	80.0	859	2	BF532181
C 81	16	80.0	903	2	BE961035
C 82	16	80.0	1420	2	BF128220
C 83	16	80.0	1461	4	BG844488
C 84	16	80.0	1974	9	AG187377
C 85	15.8	79.0	139	4	BG299044
C 86	15.8	79.0	172	8	AQ646084
C 87	15.8	79.0	173	6	BY566602
C 88	15.8	79.0	176	2	AW292530
C 89	15.8	79.0	188	1	AI206251
C 90	15.8	79.0	206	8	AQ639747
C 91	15.8	79.0	216	2	BE595788
C 92	15.8	79.0	225	8	AQ949378
C 93	15.8	79.0	236	2	BF221719
C 94	15.8	79.0	244	5	BQ042798
C 95	15.8	79.0	257	4	EM318712
C 96	15.8	79.0	273	2	AW611476
C 97	15.8	79.0	274	1	AI919398

BI833379	603088013
CK017551	AGENCOURT
CD754743	AGENCOURT
CR811620	GR0AAA35C
CD252254	AGENCOURT
AG426940	Mus muscu
BM048685	603628283
BQ707308	AGENCOURT
BI414634	602989764
CF106699	PI5P19 Un
AU076753	AU076753
BM011282	603635446
CR546210	DKF2P459A
BG408744	gb80006.Y
AW281714	fj53b11.x
BH111553	RPI-24-2
CD589669	RK053A3F0
AA067552	26183 Lam
CA041772	ssalplnb5
BM089797	503647 MA
CB504055	ssalplnb5
CA041823	ssalplnb5
CA061753	ssalxgb51
CB505188	ssalnges0
AL847928	AL847928
CA379658	658790 NC
BJ594673	BJ594673
BJ590911	BJ590911
AJ425708	AJ425708
BJ587908	BJ587908
BJ579066	BJ579066
BJ610428	BJ610428
BP142881	BP142881
CA362363	636481 NC
AG402582	Mus muscu
AL259447	Tetraodon
CR080093	Reverse s
BP170401	BP170401
AL232373	Tetraodon
CNI163617	953145 MA
AL169724	Tetraodon
AL245968	Tetraodon
BF537884	602049323
BG293564	602390214
AG574348	Mus muscu
AG321235	Mus muscu
BG4848292	102402180
BF342971	602017076
AZ296490	RPI-23-1
CR224707	Forward s
AZ625316	1M0464A16
AQ680838	HS 5488.A
AZ427072	1M0208P13
BE569899	601332108
BF532181	602073062
BE961035	601648406
BF128220	601810451
BG844488	102400680
AG187377	Pan trogl
BG299044	602397714
AQ646084	RPC193-Dp
BY566602	BY566602
AW292530	UI-H-B12-
AI206251	qz2f03.x
AQ639747	z27P1-19G
BE595788	Pil 54_A0
AQ949378	Sheared D
BF221719	7141g02.x
BQ042798	UI-N-EMO-
EM318712	Pil 16_G0
AW611476	hg91a06.x
AI919398	tm83c12.x

98	15.8	79.0	280	8	AQ661130	Sheared D	c 171	15.8	79.0	527	9	CNS03JFD	AL246802 Tetradon
99	15.8	79.0	295	9	CE037312	tigr-gss-	172	15.8	79.0	528	9	TA393G11P	AL498746 T. brucei
100	15.8	79.0	307	5	BQ291932	PM3-AN009	c 173	15.8	79.0	529	8	AQ940613	Sheared D
101	15.8	79.0	307	5	BY465526	BY465526	c 174	15.8	79.0	530	8	AQ884727	HS_5517_A
102	15.8	79.0	320	2	BB586557	BB586557	175	15.8	79.0	530	8	AQ661125	Sheared D
103	15.8	79.0	329	8	AZ592777	IM0403N21	176	15.8	79.0	533	8	AQ661125	T. brucei
104	15.8	79.0	331	2	AW139807	UI-H-B11-	c 177	15.8	79.0	541	8	AQ784412	HS 3141_A
105	15.8	79.0	335	1	AA464973	zx80g11-g	c 178	15.8	79.0	542	8	AQ650723	Sheared D
106	15.8	79.0	350	2	AW409211	fb2_c2_Fe	c 179	15.8	79.0	548	8	BZ859355	CH240_232
107	15.8	79.0	352	1	AJ397018	AJ397018	c 180	15.8	79.0	550	5	BX864609	BX864609
108	15.8	79.0	355	1	AL657936	AL657936	c 181	15.8	79.0	551	8	AQ948012	Sheared D
109	15.8	79.0	361	2	BE682601	180759_MA	c 182	15.8	79.0	553	8	AQ618520	HS 5170_B
110	15.8	79.0	367	9	TA108A02P	T. brucei	c 183	15.8	79.0	554	4	BI398572	MI-P-AV1-
111	15.8	79.0	372	8	AQ640322	927P1-2H1	184	15.8	79.0	554	7	CO067284	Mdfw2064m
112	15.8	79.0	374	8	BZ882695	CH240_241	185	15.8	79.0	554	7	TA272E10Q	T. brucei
113	15.8	79.0	375	8	AZ399414	IM0165E17	186	15.8	79.0	555	8	AQ944550	Sheared D
114	15.8	79.0	379	9	TA153H09P	TA153H09P	c 187	15.8	79.0	556	9	TA393G11Q	T. brucei
115	15.8	79.0	399	1	AI347122	qp55d05.x	c 188	15.8	79.0	561	4	BI210237	EST528277
116	15.8	79.0	402	8	AQ085410	HS 2164_B	c 189	15.8	79.0	563	2	BE857086	7923h07.x
117	15.8	79.0	402	8	AQ684123	RPCI93-Dp	c 190	15.8	79.0	563	7	CN544769	UI-R-DY1-
118	15.8	79.0	407	9	CB622473	tigr-gss-	c 191	15.8	79.0	564	2	BF078301	T. brucei
119	15.8	79.0	411	1	AI1992007	w843f12.x	c 192	15.8	79.0	566	9	TA153H09Q	T. brucei
120	15.8	79.0	411	2	BE367555	P11_9_C08	c 193	15.8	79.0	567	6	CA803739	ESG0110b.
121	15.8	79.0	411	8	AQ023239	HS 2179_B	c 194	15.8	79.0	569	5	BM887972	TMT188_Hu
122	15.8	79.0	417	7	CO695732	DG11-8911	195	15.8	79.0	569	5	BI203864	EST521904
123	15.8	79.0	420	5	BY288753	BY288753	196	15.8	79.0	571	4	AZ220932	Sheared D
124	15.8	79.0	421	1	AI768063	w146c11.x	197	15.8	79.0	573	2	BF198956	Sheared D
125	15.8	79.0	423	8	AQ948675	Sheared D	c 198	15.8	79.0	575	2	BF198956	Sheared D
126	15.8	79.0	426	8	AQ023204	AQ023204	199	15.8	79.0	575	6	CB497894	omykrbna5
127	15.8	79.0	428	9	TA156G10P	TA156G10P	200	15.8	79.0	578	4	BF997918	MR3-GNO15
128	15.8	79.0	431	1	AU088755	AU088755	c 201	15.8	79.0	579	6	CA378490	6573h07_NC
129	15.8	79.0	431	5	BY218755	BY218755	c 202	15.8	79.0	580	5	BU877392	V033E03_P
130	15.8	79.0	431	5	AQ637907	927P1-5G1	c 203	15.8	79.0	582	3	CNS09H1A	Single re
131	15.8	79.0	440	2	BE363891	P11_10_F0	c 204	15.8	79.0	586	6	CA804153	ESG0112b.
132	15.8	79.0	443	6	CB789266	AMGNNUC:S	c 205	15.8	79.0	588	7	CK132617	RSQ3692.3
133	15.8	79.0	444	8	BH057373	RPCI-24-2	206	15.8	79.0	589	5	BQ580036	ACB33-Jm8
134	15.8	79.0	445	7	CN955640	3139_102-	207	15.8	79.0	589	8	BZ896254	NARp7_018
135	15.8	79.0	449	5	EX091000	EX091000	208	15.8	79.0	608	5	BX878642	BX878642
136	15.8	79.0	449	8	BZ932053	CH240_50M	209	15.8	79.0	608	5	CA485093	CA485093
137	15.8	79.0	453	8	BH282546	BH282546	210	15.8	79.0	610	9	CL388659	RPCI44_28
138	15.8	79.0	454	9	CQ098462	Forward s	c 211	15.8	79.0	611	5	BU313895	603543071
139	15.8	79.0	455	8	AQ947640	Sheared D	212	15.8	79.0	611	5	BU344246	603523547
140	15.8	79.0	456	8	AZ236621	RPCI-23-5	213	15.8	79.0	616	9	CE698153	tigr-gss-
141	15.8	79.0	462	1	AI262680	qk35b06.x	214	15.8	79.0	617	8	BH374955	AG-ND-104
142	15.8	79.0	465	6	BY556752	BY556752	c 215	15.8	79.0	617	9	TA58D12P	T. brucei
143	15.8	79.0	467	6	CB714342	AMGNNUC:N	c 216	15.8	79.0	621	9	CL706046	PHRC-GT-
144	15.8	79.0	467	8	AQ940637	Sheared D	c 217	15.8	79.0	632	6	CA047810	ssalbrh01
145	15.8	79.0	467	8	AQ653866	Sheared D	218	15.8	79.0	636	2	BB626318	BB626318
146	15.8	79.0	467	9	TA58D12Q	T. brucei	219	15.8	79.0	636	8	AQ661405	Sheared D
147	15.8	79.0	468	2	BE502327	hy22a04.x	220	15.8	79.0	637	4	BG385464	602454347
148	15.8	79.0	473	8	AQ066534	RPCI-23-4	221	15.8	79.0	637	5	BQ771765	UI-H-E21-
149	15.8	79.0	479	5	BU394802	603806077	c 222	15.8	79.0	638	8	AQ659749	Sheared D
150	15.8	79.0	480	9	CQ931039	MBFPH94TR	223	15.8	79.0	639	4	BF999529	MR3-GNO15
151	15.8	79.0	480	9	AZ212338	Sheared D	224	15.8	79.0	646	9	CR813888	GROAA38D
152	15.8	79.0	483	1	AI861847	wal3f02.x	225	15.8	79.0	654	2	AW745415	WS1_34_E1
153	15.8	79.0	483	7	CP433978	N1F1_31_B	c 226	15.8	79.0	654	6	CA373116	647148_NC
154	15.8	79.0	484	8	AQ762564	HS_191_B	c 227	15.8	79.0	656	1	AI758825	ty24h06.x
155	15.8	79.0	490	8	AQ944554	Sheared D	c 228	15.8	79.0	656	5	BU292237	604164707
156	15.8	79.0	492	8	BH035715	RPCI-24-3	c 229	15.8	79.0	656	9	CE159452	tigr-gss-
157	15.8	79.0	493	8	AQ947021	Sheared D	c 230	15.8	79.0	660	6	CA386599	68228_NC
158	15.8	79.0	494	1	AI698016	we19h05.x	c 231	15.8	79.0	662	8	AQ653867	Sheared D
159	15.8	79.0	496	8	AQ947123	Sheared D	c 232	15.8	79.0	668	4	BJ736737	BJ736737
160	15.8	79.0	500	1	AI367097	qk6e01.x	c 233	15.8	79.0	670	5	BQ446670	UI-H-EU1-
161	15.8	79.0	502	8	AQ647248	RPCI93-Dp	c 234	15.8	79.0	675	6	CA361364	635170_NC
162	15.8	79.0	506	1	AL675840	AL675840	235	15.8	79.0	676	8	BZ121026	CH230-366
163	15.8	79.0	510	7	CK743721	eca01-12C	236	15.8	79.0	676	9	CE613987	tigr-gss-
164	15.8	79.0	512	6	CA502614	WHB4302_F	c 237	15.8	79.0	678	8	AQ162760	mgxb0016A
165	15.8	79.0	515	9	AQ763683	HS 2162_B	238	15.8	79.0	679	5	BQ772276	UI-H-E21-
166	15.8	79.0	515	9	CB341227	RPCI44_26	239	15.8	79.0	686	4	BI454025	603174926
167	15.8	79.0	521	8	AQ659748	Sheared D	240	15.8	79.0	694	7	CN113142	ia94f09.9
168	15.8	79.0	522	8	BZ38341	ia94f09.b	c 241	15.8	79.0	694	7	CN113142	EC2CAA39D
169	15.8	79.0	522	8	AQ650766	Sheared D	242	15.8	79.0	700	6	BY750526	BY750526
170	15.8	79.0	525	9	TA91A08P	TA91A08P	243	15.8	79.0	701	4	BG470159	602533861

[illegible]

C 390	15.4	77.0	520	2	AW018661	AW018661 fd59f08.y	C 463	15.4	77.0	771	7	CF348185	CF348185 AGENCOURT
C 391	15.4	77.0	521	4	BM592892	BM592892 170006874	C 464	15.4	77.0	773	7	CK239244	CK239244 AGENCOURT
C 392	15.4	77.0	523	4	BI473314	BI473314 fd42c11.y	C 465	15.4	77.0	773	7	CO798628	CO798628 AGENCOURT
C 393	15.4	77.0	526	8	A2411805	A2411805 1M0193E12	C 466	15.4	77.0	774	8	BZ801690	BZ801690 PUFN05TD
C 394	15.4	77.0	527	6	CD479126	CD479126 ecs01-37m	C 467	15.4	77.0	775	4	BI734895	BI734895 603356094
C 395	15.4	77.0	528	2	AW077355	AW077355 fj34b07.y	C 468	15.4	77.0	777	7	CK024292	CK024292 AGENCOURT
C 396	15.4	77.0	528	2	AW077356	AW077356 fj34b08.y	C 469	15.4	77.0	777	7	CN503406	CN503406 AGENCOURT
C 397	15.4	77.0	536	4	BM035913	BM035913 fu21h08.y	C 470	15.4	77.0	778	7	CN513246	CN513246 AGENCOURT
C 398	15.4	77.0	537	1	AI384791	AI384791 fb17d06.y	C 471	15.4	77.0	779	9	CN501984	CN501984 AGENCOURT
C 399	15.4	77.0	539	4	BI708671	BI708671 fq11e06.y	C 472	15.4	77.0	779	9	AG595594	AG595594 Mus muscu
C 400	15.4	77.0	541	1	AI486352	AI486352 EST244673	C 473	15.4	77.0	780	8	BZ801689	BZ801689 PUFN05TB
C 401	15.4	77.0	543	6	CD283886	CD283886 G39167.54	C 474	15.4	77.0	781	7	CO958346	CO958346 AGENCOURT
C 402	15.4	77.0	553	6	CD284102	CD284102 G39172.11	C 475	15.4	77.0	782	7	CO795851	CO795851 AGENCOURT
C 403	15.4	77.0	557	8	A0769251	A0769251 HS_3222.A	C 476	15.4	77.0	783	9	CC915466	CC915466 t087j18ba
C 404	15.4	77.0	557	9	CE340465	CE340465 tigr-gss-	C 477	15.4	77.0	785	7	CO9113906	CO9113906 AGENCOURT
C 405	15.4	77.0	558	8	BH391421	BH391421 AG-ND-178	C 478	15.4	77.0	788	6	CB996681	CB996681 AGENCOURT
C 406	15.4	77.0	565	5	BQ263396	BQ263396 faa22d12.r	C 479	15.4	77.0	792	4	BG620356	BG620356 604618665
C 407	15.4	77.0	567	9	BX126187	BX126187 Danio rer	C 480	15.4	77.0	792	7	CK363246	CK363246 AGENCOURT
C 408	15.4	77.0	572	4	BJ486927	BJ486927 BJ486927	C 481	15.4	77.0	794	7	CK018529	CK018529 AGENCOURT
C 409	15.4	77.0	573	8	A2417144	A2417144 1M0192M11	C 482	15.4	77.0	797	7	CK238489	CK238489 AGENCOURT
C 410	15.4	77.0	574	8	AQ149649	AQ149649 HS_3171.B	C 483	15.4	77.0	797	9	CC577857	CC577857 CH240.456
C 411	15.4	77.0	593	2	AW258188	AW258188 uq31e09.y	C 484	15.4	77.0	802	7	CN500782	CN500782 AGENCOURT
C 412	15.4	77.0	596	8	AZ397663	AZ397663 1M0162E18	C 485	15.4	77.0	803	4	BI891187	BI891187 ZF637-3-0
C 413	15.4	77.0	598	2	BB611350	BB611350 BB611350	C 486	15.4	77.0	806	7	CK363035	CK363035 AGENCOURT
C 414	15.4	77.0	604	5	BX952796	BX952796 DKFZp781H	C 487	15.4	77.0	807	7	CK360886	CK360886 AGENCOURT
C 415	15.4	77.0	610	9	CE504946	CE504946 tigr-gss-	C 488	15.4	77.0	810	7	CK032021	CK032021 AGENCOURT
C 416	15.4	77.0	611	6	CD284028	CD284028 G39172.73	C 489	15.4	77.0	811	7	CO956973	CO956973 AGENCOURT
C 417	15.4	77.0	614	8	AZ083342	AZ083342 RPCI-23-4	C 490	15.4	77.0	812	7	CK353203	CK353203 AGENCOURT
C 418	15.4	77.0	616	4	BG080171	BG080171 H3050E11-	C 491	15.4	77.0	814	7	CK127851	CK127851 AGENCOURT
C 419	15.4	77.0	617	5	BX952802	BX952802 DKFZp781I	C 492	15.4	77.0	814	7	CK363292	CK363292 AGENCOURT
C 420	15.4	77.0	623	8	AZ595199	AZ595199 1M0407I12	C 493	15.4	77.0	822	7	CF549869	CF549869 AGENCOURT
C 421	15.4	77.0	623	9	DR25J135	DR25J135 Danio rer	C 494	15.4	77.0	824	7	CK240309	CK240309 AGENCOURT
C 422	15.4	77.0	626	9	LBAP046E04	LBAP046E04 Leishmani	C 495	15.4	77.0	830	7	CK357342	CK357342 AGENCOURT
C 423	15.4	77.0	629	6	CD284464	CD284464 G39177.3	C 496	15.4	77.0	836	7	CN177917	CN177917 AGENCOURT
C 424	15.4	77.0	630	9	CW519936	CW519936 OP_Ba002	C 497	15.4	77.0	839	7	CN326110	CN326110 AGENCOURT
C 425	15.4	77.0	633	8	BZ846168	BZ846168 CH240.252	C 498	15.4	77.0	840	6	CD575446	CD575446 UCPT01.0
C 426	15.4	77.0	634	4	BM095467	BM095467 fv31c02.y	C 499	15.4	77.0	847	7	CK237209	CK237209 AGENCOURT
C 427	15.4	77.0	641	5	BQ261141	BQ261141 fz78a08.y	C 500	15.4	77.0	848	7	CO561198	CO561198 AGENCOURT
C 428	15.4	77.0	654	8	B65483	B65483 CIT-HSP-202							
C 429	15.4	77.0	659	9	CNS01GT4	AL143561 Anopheles							
C 430	15.4	77.0	686	2	BF101951	BF101951 601752917							
C 431	15.4	77.0	687	7	CF489298	CF489298 POLI_56.A							
C 432	15.4	77.0	689	7	CN021018	CN021018 AGENCOURT							
C 433	15.4	77.0	694	6	CB505877	CB505877 ssalp1nb5							
C 434	15.4	77.0	695	1	AL821998	AL821998 AL821998							
C 435	15.4	77.0	703	4	BI931176	BI931176 EST551065							
C 436	15.4	77.0	703	9	CR117435	CR117435 Forward s							
C 437	15.4	77.0	704	9	CC830687	CC830687 ZMMBB018							
C 438	15.4	77.0	707	9	AG348156	AG348156 Mus muscu							
C 439	15.4	77.0	718	9	AG216491	AG216491 Drosophil							
C 440	15.4	77.0	721	7	CV118728	CV118728 AGENCOURT							
C 441	15.4	77.0	721	9	AG577272	AG577272 Mus muscu							
C 442	15.4	77.0	721	9	CU169457	CU169457 104_368.1							
C 443	15.4	77.0	723	9	AG400232	AG400232 Mus muscu							
C 444	15.4	77.0	725	8	BZ805405	BZ805405 PUGFBA8TB							
C 445	15.4	77.0	730	7	CN508667	CN508667 AGENCOURT							
C 446	15.4	77.0	733	7	CN502981	CN502981 AGENCOURT							
C 447	15.4	77.0	733	7	CO926376	CO926376 AGENCOURT							
C 448	15.4	77.0	745	7	CO796922	CO796922 AGENCOURT							
C 449	15.4	77.0	747	6	CA318484	CA318484 UI-M-FW0-							
C 450	15.4	77.0	748	6	CD803103	CD803103 UI-M-GV0-							
C 451	15.4	77.0	749	9	CU169458	CU169458 104_368.1							
C 452	15.4	77.0	750	7	CO925439	CO925439 AGENCOURT							
C 453	15.4	77.0	756	7	CK027436	CK027436 AGENCOURT							
C 454	15.4	77.0	758	9	AG608895	AG608895 Mus muscu							
C 455	15.4	77.0	759	7	CN322633	CN322633 AGENCOURT							
C 456	15.4	77.0	760	7	CO920588	CO920588 AGENCOURT							
C 457	15.4	77.0	761	9	AG553135	AG553135 Mus muscu							
C 458	15.4	77.0	762	7	CN326173	CN326173 AGENCOURT							
C 459	15.4	77.0	763	7	CK236088	CK236088 AGENCOURT							
C 460	15.4	77.0	764	7	CN014892	CN014892 AGENCOURT							
C 461	15.4	77.0	764	7	CN018560	CN018560 AGENCOURT							
C 462	15.4	77.0	767	7	CK018730	CK018730 AGENCOURT							

ALIGNMENTS

RESULT 1	CD523014/c	
LOCUS	CD523014	220 bp mRNA linear EST 06-JUN-2003
DEFINITION	AGENCOURT 14358397 NIH_MGC_191 Homo sapiens cDNA clone IMAGE:30410101 5', mRNA sequence.	
ACCESSION	CD523014	
VERSION	CD523014.1	GI:31454792
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 220)	
TITLE	NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-r@mail.nih.gov Tissue Procurement: Narayan Bhat cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDCM199 row: 1 column: 14 High quality sequence stop: 105.	

FEATURES
source

Location/Qualifiers
1..220
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410101"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site 1: Sfil (ggccattatggcc); Site 2: Sfil (ggcgctcgcc); Library is oligo-dT primed and directionally cloned. PBMC - peripheral blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA and Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 61 CGTGTGCTGTGCTAGTCCC 42

RESULT 2

LOCUS CD698408/0 490 bp mRNA linear EST 25-JUN-2003
DEFINITION EST14931 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD698408
VERSION CD698408.1 GI:32226772
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and Zeng, Y.-X.

TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: YiXin Zeng
Cancer Center

Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@zsu.edu.cn.

FEATURES
source

Location/Qualifiers
1..490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 490;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||

Db

97 CGTGTGCTGTGCTAGTCCC 78

RESULT 3

LOCUS BX371662 626 bp mRNA linear EST 23-APR-2004
DEFINITION BX371662 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI027YM22 3-PRIME, mRNA sequence.
ACCESSION BX371662
VERSION BX371662.2 GI:46558109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 626)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30437975.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2188.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BA1020ZB01_CS01826_1&c=2188.f

FEATURES

source

Location/Qualifiers
1..626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI027YM22"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 626;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 548 CGTGTGCTGTGCTAGTCCC 567

RESULT 4

LOCUS BQ109523 690 bp mRNA linear EST 16-APR-2002
DEFINITION imageqc7_2001/bnm347bdr81.y1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203544 5', mRNA sequence.
ACCESSION BQ109523
VERSION BQ109523.1 GI:20159177
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 690)
AUTHORS Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.

TITLE The I.M.A.G.E. Consortium quality control effort: clone resequencing for verification

JOURNAL
COMMENT

Unpublished (2001)
Other ESTs: BI767024
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov

This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.

Plate: LHAM11510 row: d column: 9

Seq primer: ml3rpl

High quality sequence stop: 690.

FEATURES

source

1. .690

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5203544"

/lab_host="DH10B"

/clone_lib="NIH_MGC_122"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 690;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTATCCC 20

Db 64 CGTGTGCTGTGCTAGTATCCC 45

RESULT 5

BI767024/c

LOCUS

DEFINITION BI767024 695 bp mRNA linear EST 25-SEP-2001 603054234F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203544 5', mRNA sequence.

ACCESSION BI767024.1 GI:15758602

VERSION 1 (bases 1 to 695)

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM11510 row: d column: 09

FEATURES

source

High quality sequence stop: 637.

Location/Qualifiers

1. .695

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5203544"

/lab_host="DH10B"

/clone_lib="NIH_MGC_122"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 695;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTATCCC 20

Db 65 CGTGTGCTGTGCTAGTATCCC 46

RESULT 6

BI0109553/c

LOCUS

DEFINITION BI0109553 709 bp mRNA linear EST 16-APR-2002 imageqc_7 2001/anns58bdr81.y1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5218562 5', mRNA sequence.

ACCESSION BI0109553

VERSION BQ109553.1 GI:20159207

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 709)

Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and

Prange, C.K.

The I.M.A.G.E. Consortium quality control effort: clone

resequencing for verification

Unpublished (2001)

Contact: Prange CK

The I.M.A.G.E. Consortium

Lawrence Livermore National Laboratory

Livermore, CA, USA

Email: help@image.llnl.gov

This read has been produced as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.

Plate: LHAM11549 row: f column: 3

Seq primer: ml3rpl

High quality sequence stop: 709.

FEATURES

source

1. .709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5218562"

/tissue_type="leukocyte"

/lab_host="DH10B"

/clone_lib="NIH_MGC_118"

/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 709;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 20
|||||
Db 63 CGTGTGCTGTGCTAGTCC 44

RESULT 7

CD691973/c
LOCUS 132 bp mRNA linear EST 25-JUN-2003
DEFINITION EST8496 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD691973
VERSION CD691973.1 GI:32214208
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 132)
Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
Zeng, Y.-X.
Transcriptional Gene Expression Profile of Human Nasopharynx

TITLE

JOURNAL

COMMENT

Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.

FEATURES

source

1..132
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_libs="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 95.0%; Score 19; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 19
|||||
Db 99 CGTGTGCTGTGCTAGTCC 81

RESULT 8

BI906246/c
LOCUS 753 bp mRNA linear EST 16-OCT-2001
DEFINITION 603063172F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212648 5',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 753)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11533 row: 0 column: 17
High quality sequence stop: 719.

FEATURES

source

1..753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5212648"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 95.0%; Score 19; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 19
|||||
Db 56 CGTGTGCTGTGCTAGTCC 38

RESULT 9

BI824940/c
LOCUS 655 bp mRNA linear EST 04-OCT-2001
DEFINITION 603032554F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173789 5',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 655)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM11432 row: 1 column: 14
High quality sequence start: 27
High quality sequence stop: 653.

```

FEATURES
source
Location/Qualifiers
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173789"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
sample anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 4; Length 655;
Best Local Similarity 95.0%; Pred. No. 3e+02; 1; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 20
|||||
Db 98 CGTGTGCTGTGCTAGTCC 79
|||||

RESULT 10
CO647434/c
LOCUS
DEFINITION
ILLUMIGEN MCQ 40481 Katze WMPB2 Macaca mulatta cDNA clone
IBUW:23771 5' similar to Bases 130 to 980 highly similar to human
CD86 (Hs:27954), mRNA sequence.
ACCESSION
CO647434
VERSION
CO647434.1 GI:50568928
KEYWORDS
EST.
SOURCE
Macaca mulatta (rhesus monkey)
ORGANISM
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE
1 (bases 1 to 995)
Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magnus,C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
JOURNAL
Contact: C. Magnus
ILLUMIGEN Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.05.28. 775 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
ILLUMIGEN Biosciences Inc. For further information, see
http://www.macaque.org
PCR Primers
FORWARD: CCTCACTAAAGGACAAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 995 Std Error: 0.00
Plate: CL000337 row: C column: 06
Seq primer: CCTCACTAAAGGACAAAA
POLYA=yes.
Location/Qualifiers
1. .995
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBUW:23771"
/sex="male"
/tissue_type="blood"

```

```

/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_lib="Katze WMPB2"
/notes="Vector: pDONR 222; Site 1: BsrG I; Site 2: BsrG I;
Created from CloneMiner cDNA Library Construction kit
(catalog #18249-029)"

ORIGIN
Query Match 92.0%; Score 18.4; DB 7; Length 995;
Best Local Similarity 95.0%; Pred. No. 3.1e+02; 1; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 67 CGTGTGCTGTGCTAGTCCC 48
|||||

RESULT 11
AI992726
LOCUS
DEFINITION
701493407 A. thaliana, Ohio State clone set arabidopsis thaliana
cDNA clone 701493407, mRNA sequence.
ACCESSION
AI992726
VERSION
AI992726.1 GI:5839631
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 542)
Chen,J., Moniyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D.,
Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C.,
Carpio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S.,
Nobriga,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and
Hanson,D.
Arabidopsis thaliana Gene Expression MicroArray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
Location/Qualifiers
1. .542
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="701493407"
/clone_lib="A. thaliana, Ohio State clone set"
/notes="cDNA library was made from selected clones from the
Arabidopsis thaliana Ohio State clone set."

ORIGIN
Query Match 87.0%; Score 17.4; DB 1; Length 542;
Best Local Similarity 94.7%; Pred. No. 9.1e+02; 1; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGTGCTAGTCCC 20
|||||
Db 94 GTGTGTCTGTGCTATCCC 112
|||||

RESULT 12
CC488528/c
LOCUS
DEFINITION
CH240_320F23.T7 CHORI-240 Bos taurus genomic clone CH240_320F23,
genomic survey sequence.

```

```

ACCESSION CC488528
VERSION CC488528.1 GI:31799360
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Holt,R., Stott,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M., Chiu,R.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dallrymple,B.P. and Tellam,R.
TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL Unpublished (2003)
COMMENT Other GSSs: CH240_320F23.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 320 row: F column: 23
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..725
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_320F23"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"
ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 725;
Best Local Similarity 94.7%; Pred. No. 9.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCC 19
|||||
Db 445 CGTGTGCTGTGCTAGCCC 427

RESULT 13
BZ166194 864 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-463117.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-463117, genomic survey sequence.
ACCESSION BZ166194
VERSION BZ166194.1 GI:23807245
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

CC488528.1 (bases 1 to 864)
Zhao,S., Shetty,J., Shatsman,S., Tsagay,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,P., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-463117.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). BAC end
plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 463 row: I column: 17
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..864
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHad/MCW"
/db_xref="taxon:10116"
/clone="CH230-463117"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SSHad/MCW) BAC library produced by
Pieter de Jong"
ORIGIN
Query Match 87.0%; Score 17.4; DB 8; Length 864;
Best Local Similarity 94.7%; Pred. No. 9.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGTCTGTGCTAGTCCC 20
|||||
Db 820 GTGTGTCTGTGCTAGTCC 838

RESULT 14
BF304344/c 1126 bp mRNA linear EST 21-NOV-2000
LOCUS 601887245F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4121195 5',
DEFINITION mRNA sequence.
ACCESSION BF304344
VERSION BF304344.1 GI:11251069
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM1002 row: j column: 12
High quality sequence stop: 638.

```

```

FEATURES
  source
    Location/Qualifiers
      1. 1126
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4121195"
        /tissue_type="rhabdomyosarcoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_17"
        /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
        Site 2: XhoI; cDNA made by oligo-dT priming.
        Directionally cloned into EcoRI/XhoI sites using the
        following 5' adaptor: GGCACGAG(G). Size-selected >500bp
        for average insert size 1.8kb. Library constructed by
        Ling Hong in the laboratory of Gerald M. Rubin (University
        of California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

  Query Match      87.0%; Score 17.4; DB 2; Length 1126;
  Best Local Similarity 94.7%; Pred. No. 9.7e+02;
  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2 GTGTGTCGTCTGCTAGTCCC 20
    |||||
Db  889 GTGCTCTGTGCTAGTCCC 871

RESULT 15
CL508080/c
LOCUS
DEFINITION
  CL508080      2004 bp      DNA      linear      GSS 01-APR-2004
  SAIL_792_D02_v1 SAIL Collection Arabidopsis thaliana genomic clone
  SAIL_792_D02_v1, genomic survey sequence.
ACCESSION
  CL508080
VERSION
  CL508080.1 GI:46005400
KEYWORDS
  GSS.
SOURCE
  Arabidopsis thaliana (chale cress)
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
  1 (bases 1 to 2004)
REFERENCE
  Sessions A., Burke E., Presting G., Aux G., McElver J., Patton D.,
  Dietrich B., Ho P., Bacwaden J., Ko C., Clarke J.D., Cotton D.,
  Bullis D., Snell J., Miguel T., Hutchison D., Kimmerly B.,
  Mitzel T., Katagiri F., Glazebrook J., Law M. and Goff S.A.
  A high-throughput Arabidopsis reverse genetics system
  Plant Cell 14 (12), 2985-2994 (2002)
  22356987
  Plant Cell 14 (12), 2985-2994 (2002)
  12468722
  Contact: Sessions A
  Applied Trait Genetics
  Syngenta Biotechnology Inc.
  3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
  Email: allen.sessions@syngenta.com
  ABRC Stock Number CS835414; T-DNA left border flanking sequences of
  Syngenta Arabidopsis Insertion Library (SAIL) lines are available
  through the Arabidopsis Biological Resource Center (ABRC).
  Sequences represent a pool of amplified genomic regions and not
  single contiguous sequences.
  Class: T-DNA tagged.
  Location/Qualifiers
    1. 2004
      /organism="Arabidopsis thaliana"
      /mol_type="genomic DNA"
      /ecotype="Columbia"
      /db_xref="taxon:3702"
      /clone="SAIL_792_D02_v1"
      /clone_lib="SAIL_Collection"
      /note="T-DNA left border sequences were isolated using a
      modified TAIL-PCR strategy"
```

ORIGIN

```

FEATURES
  source
    Location/Qualifiers
      1. 1245
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6169502"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_72"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 2 Kb. Library constructed by Life
        Technologies."
```

ORIGIN

```

  Query Match      85.0%; Score 17; DB 5; Length 1245;
  Best Local Similarity 100.0%; Pred. No. 1.5e+03;
  Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 CGTGTGTCGTCTGCTAGTCC 17
    |||||
Db  1209 CGTGTGTCGTCTAGT 1225

RESULT 17
BY670167
LOCUS
DEFINITION
  BY670167      363 bp      mRNA      linear      EST 16-DEC-2002
  RIKEN full-length enriched, 14.5 days embryo df/df
  Rathke's pouches Mus musculus cDNA clone K82011G18 3', mRNA
  sequence.
ACCESSION
  BY670167
VERSION
  BY670167.1 GI:27050562
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Qy  1 CGTGTGTCGTCTGCTAGTCC 19
    |||||
Db  1321 CGTGTGTCGTCTAGTCC 1303

RESULT 16
BU195329
LOCUS
DEFINITION
  BU195329      1245 bp      mRNA      linear      EST 04-SEP-2002
  AGENCOURT 7964738 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6169502
  5', mRNA sequence.
ACCESSION
  BU195329
VERSION
  BU195329.1 GI:22709313
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1245)
REFERENCE
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: rsgrbs@mail.nih.gov
  Tissue Procurement: ATCC/DCTD/DTF
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LMNL at:
  http://image.llnl.gov
  Plate: LHAM13534 row: 1 column: 15
  High quality sequence stop: 233.
  Location/Qualifiers
    1. 1245
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6169502"
      /tissue_type="melanotic melanoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_72"
      /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
      Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
      Average insert size 2 Kb. Library constructed by Life
      Technologies."
```

1 (bases 1 to 363)
 Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
 Nikaide,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
 Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
 Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
 Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Mateuda,H.,
 Batalov,S., Beisel,K.W., Blake,J.A., Bratt,D., Brusic,V.,
 Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
 Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
 Gariboldi,M., Glasi,C., Godzik,A., Gough,J., Grimmond,S.,
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
 Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konggaya,A.,
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
 Numata,K., Okido,T., Pavan,W.J., Perteza,G., Pesole,G.,
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
 Sandelin,A., Schneider,C., Sempke,C.A., Setou,M., Shimada,K.,
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
 Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Sushiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
 Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
 Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
 Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
 Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
 Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
 Hayashizaki,Y.: Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Michelle Brinkmeier and Sally Camper (
 Dept. Human Genetics University of Michigan Medical School 4301
 MSRB 3 1500 W. Medical Center Dr. Ann Arbor,MI 48109-0638 USA)
 whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 Location/Qualifiers
 1. .363
 /organism="Mus musculus"
 /mol_type="mRNA"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 6; Length 363;
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGTGTGCTGTGCTAGTATCCC 20
 DB 59 CGTGTGCTGTGCTAGTATCCC 78
 RESULT 18
 BX963297 395 bp DNA linear GSS 05-JUL-2004
 LOCUS Reverse strand read from insert in 5'HPRT insertion targeting and
 DEFINITION chromosome engineering clone MHPN136115, genomic survey sequence.
 ACCESSION BX963297
 VERSION BX963297.1 GI:49694720
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 395)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
 FEATURES Location/Qualifiers
 source 1. .395
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPN136115"
 /clone_lib="MHPN"
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 395;
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGTGTGCTGTGCTAGTATCCC 20
 DB 305 CGTGTGCTGTGCTAGTATCCC 324
 RESULT 19
 CF805725 571 bp mRNA linear EST 15-APR-2004
 LOCUS pSHA0071a11r Agriculture Canada Phytoththora sojae EST project
 DEFINITION Glycine max cDNA clone SHA007A11 5, mRNA sequence.
 ACCESSION CF805725
 VERSION CF805725.1 GI:37993979
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 571)
 AUTHORS Qutob,D., Hraber,P.T., Sobral,B.W.S. and Gijzen,M.
 TITLE Comparative analysis of expressed sequences in Phytophthora sojae
 JOURNAL Plant Physiol. 123 (1), 243-254 (2000)
 MEDLINE 20267956

```

PUBMED      10806241
COMMENT      Contact: Tyler B
              Tyler lab
              VBI
              1880 Pratt Dr., Blacksburg, VA 24061, USA
              Tel: 540-231-7318
              Email: bmtyle@vt.edu
              PCR Primers
              FORWARD: BK reverse primer
              BACKWARD: BK reverse primer
              Plate: 007 row: A column: 11
              Seq primer: BK reverse primer
              High quality sequence stop: 571.

FEATURES     source
              1..571
              /organism="Glycine max"
              /mol_type="mRNA"
              /cultivar="Harosoy"
              /db_xref="taxon:3847"
              /clone="sUA007A11"
              /tissue_type="Phytophthora sojae-infected hypocotyl"
              /cell_line="Phytophthora sojae culture P6497"
              /dev_stage="48 hr. post infection stage"
              /clone_lib="Agriculture Canada Phytophthora sojae EST
              project"
              /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 7; Length 571;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 20 CGTGTGCTGTGCTGTGCC 1

RESULT 20
LOCUS      BX968677 615 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 5'HPRT insertion targeting and
              chromosome engineering clone MHPN410b24, genomic survey sequence.
ACCESSION  BX968677
VERSION     BX968677.1 GI:49700100
KEYWORDS    GSS; Genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 615)
AUTHORS     Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
              Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
              Rogers,J. and Bradley,A.
TITLE       Direct Submision
JOURNAL     Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES     source
              1..615
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              /clone="MHPN410b24"
              /clone_lib="MHPN"

ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 615;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 179 CGTGTGCTGTGCCAGTCAC 160

```

```

RESULT 21
BZ285459/c
LOCUS      BZ285459 659 bp DNA linear GSS 15-OCT-2002
DEFINITION CH230-336K14.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
              CH230-336K14, genomic survey sequence.
ACCESSION  BZ285459
VERSION     BZ285459.1 GI:24016470
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE   1 (bases 1 to 659)
AUTHORS     Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K.,
              Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
              Riggs,F., de Jong,P. and Fraser,C.M.
TITLE       Rat BAC End Sequences from Library CHORI-230 MboI segment
              Unpublished (1999)
COMMENT      Other GSSs: CH230-336K14.TV
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/or ering information.htm). BAC end
              page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
              Plate: 336 row: K column: 14
              Seq primer: SP6
              Class: BAC ends.

FEATURES     Location/Qualifiers
              1..659
              /organism="Rattus norvegicus"
              /mol_type="genomic DNA"
              /strain="BN/SeNHsd/MCW"
              /db_xref="taxon:10116"
              /clone="CH230-336K14"
              /sex="Female"
              /cell_type="Brain"
              /clone_lib="CHORI-230 Segment 2"
              /notes="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
              CHORI-230 Rat (BN/SeNHsd/MCW) BAC library produced by
              Pieter de Jong"

ORIGIN
Query Match      84.0%; Score 16.8; DB 8; Length 659;
Best Local Similarity 90.0%; Pred. No. 1.8e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 359 CATGTGCTGTGCTAGGCC 340

RESULT 22
BZ339267/c
LOCUS      BZ339267 666 bp mRNA linear EST 31-JAN-2000
DEFINITION X289f02.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2871387 3',
              similar to gb:M22489 BONE MORPHOGENETIC PROTEIN 2 PRECURSOR
              (HUMAN); mRNA sequence.
ACCESSION  BZ339267
VERSION     BZ339267.1 GI:6835926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1 (bases 1 to 666)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbhp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 455.
 Location/Qualifiers
 1..666
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2871387"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 666;
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20

Db 574 CGTGTGCTGTGCTAGTCA 593

RESULT 23

LOCUS CR120951 696 bp DNA linear GSS 06-JUL-2004
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP74c05, genomic survey sequence.

ACCESSION CR120951

VERSION CR120951.1 GI:49868401

KEYWORDS GSS; genome survey sequence; MICER.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 696)
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

Location/Qualifiers

1..696

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHP74c05"

/clone_lib="MHPp"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 696;
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20

Db 466 CGTGGTCTGTGCTAGTCC 485

RESULT 24

LOCUS AG353788 760 bp DNA linear GSS 02-JUN-2004
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-151L04.T7, genomic survey sequence.

ACCESSION AG353788

VERSION AG353788.1 GI:47927098

KEYWORDS GSS.

SOURCE Mus musculus molossinus

ORGANISM

Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

BAC end Sequences of Library MSMg01

Unpublished

2' (bases 1 to 760)

Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

Direct Submission

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@psc.riken.jp URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-45-503-9111 Fax:81-45-503-9170)

Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute, Bio Resource Center,
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyada, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

source

1..760

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-151L04.T7"

/sex="male"

/tissue_type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 760;
 Best Local Similarity 90.0%; Pred. No. 1.8e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20

Db 460 CGTGTGCTGTGCTAATCC 479

RESULT 25

LOCUS BI833379 840 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603088013F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5227083 5',

```

mRNA sequence.
ACCESSION BI833379
VERSION BI833379.1 GI:15944929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 840)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1571 row: i column: 04
High quality sequence stop: 835.
FEATURES
Location/Qualifiers
1..840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5227083"
/lab_host="DH10B"
/clone_lib="NIH MGC 120"
/notes="Organ: pooled pancreas and spleen; Vector:
PCMV-SPT65; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 84.0%; Score 16.8; DB 4; Length 840;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGTGTGCTGTGCTAGTATCCC 20
|||||
Db 803 CGTGTGCTGTGCAAGTCTC 822
|||||

RESULT 26
CK017551/c
LOCUS CK017551 880 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT 16543712 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7044589
5', mRNA sequence.
ACCESSION CK017551
VERSION CK017551.1 GI:38543475
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 880)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen Corp
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14605 row: e column: 12
High quality sequence stop: 13
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
1..882
/organism="Danio rerio"

```

```

Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14806 row: b column: 11
High quality sequence stop: 685.
FEATURES
Location/Qualifiers
1..880
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7044589"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."
ORIGIN
Query Match 84.0%; Score 16.8; DB 7; Length 880;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGTGTGCTGTGCTAGTATCCC 20
|||||
Db 100 CGTGTGCTGTGCTGTGCTGCC 81
|||||

RESULT 27
CD754743/c
LOCUS CD754743 882 bp mRNA linear EST 30-JUN-2003
DEFINITION AGENCOURT 14618643 NCI_CGAP_Zemb2 Danio rerio cDNA clone
IMAGE:6964789 5', mRNA sequence.
ACCESSION CD754743
VERSION CD754743.1 GI:32339030
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 882)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14605 row: e column: 12
High quality sequence stop: 13
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
1..882
/organism="Danio rerio"

```

```

/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6964789"
/tissue_type="embryo"
/lab_host="DH10B (Tl-resistant)"
/clone_lib="NCI_CGAP_256B2"
/notes="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

```

ORIGIN

```

Query Match      84.0%; Score 16.8; DB 6; Length 882;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    |||||
Db 150 CGTGTGTGTGCTAGTCC 131

```

RESULT 28

CR811620/c

```

LOCUS
DEFINITION
GROAAA35CB12RM1 INRA BAC Bos taurus genomic clone INRA_597D12, DNA
sequence, genomic survey sequence.

```

ACCESSION CR811620

VERSION CR811620.1

KEYWORDS GI:52692791

SOURCE Bos taurus (cow)

ORGANISM

```

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

```

REFERENCE 1 (bases 1 to 928)

AUTHORS Eggen, A., Schibler, L. and Roy, A.

TITLE Bovine BAC End Sequences from the INRA bovine BAC library

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 928)

AUTHORS Genoscope.

TITLE Direct Submission

```

JOURNAL Submitted (20-SEP-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

```

COMMENT

```

Contact: Andre Eggen
Department of Animal Genetics - LcBC
INRA
78350 Jouy-en-Josas, France
Tel: 33 1 34 65 24 24
Fax: 33 1 34 65 24 78
Email: eggen@jouy.inra.fr

```

```

Clones are derived from the INRA bovine BAC library
(http://locus.jouy.inra.fr/fpc/cattle_bac_map.htm). For BAC library
availability, please contact Andre Eggen (eggen@jouy.inra.fr). This
work was undertaken as part of the International Bovine BAC
Mapping Consortium (IBBMC) by INRA (Jouy-en-Josas) and Genoscope
(Evry). Plate: 597 row: D column: 12
Seq primer: M13 Reverse

```

FEATURES

source

```

Location/Qualifiers
1..928

```

```

/organism="Bos taurus"
/mol_type="genomic DNA"

```

```

/strain="breed: Holstein"
/db_xref="taxon:9913"

```

```

/clone="INRA_597D12"
/sex="Male"

```

```

/cell_type="fibroblast"
/clone_lib="INRA bovine BAC"

```

```

/notes="Vector: pBeloBAC11; Site 1: HindIII; Holstein bull;
INRA Bovine BAC library (Male) produced by Andre
Eggen-Genoscope sequence ID : GROAAA35CB12RM1"

```

ORIGIN

```

Query Match      84.0%; Score 16.8; DB 9; Length 928;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    |||||
Db 546 CGTGTGTGTGCTAGTCC 527

```

RESULT 29

CD252254/c

LOCUS

```

DEFINITION
AGENCOURT 14161901 NIH_MGC 181 Homo sapiens cDNA clone
IMAGE:30374576 5', mRNA sequence.

```

ACCESSION CD252254

VERSION CD252254.1

KEYWORDS GI:31012720

SOURCE EST.

ORGANISM Homo sapiens (human)

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 937)

AUTHORS NIH-MGC

TITLE http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

```

Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

```

```

Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: NDAM434 row: d column: 09
High quality sequence stop: 16
High quality sequence stop: 508.

```

```

Location/Qualifiers
1..937

```

```

/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:30374576"

```

```

/tissue_type="White Matter"
/dev_stage="Unknown"

```

```

/lab_host="DH10B-Ton A ( T1 and T5 phage resistance)"
/clone_lib="NIH_MGC 181"

```

```

/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(deprecated); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

```

```

Query Match      84.0%; Score 16.8; DB 6; Length 937;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    |||||
Db 410 CGTGTCTGTGCTATGCC 391

```

FEATURES

source

```

Location/Qualifiers
1..937

```

```

/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:30374576"

```

```

/tissue_type="White Matter"
/dev_stage="Unknown"

```

```

/lab_host="DH10B-Ton A ( T1 and T5 phage resistance)"
/clone_lib="NIH_MGC 181"

```

```

/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(deprecated); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

```

```

Query Match      84.0%; Score 16.8; DB 6; Length 937;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    |||||
Db 410 CGTGTCTGTGCTATGCC 391

```

ORIGIN

```

Query Match      84.0%; Score 16.8; DB 6; Length 937;
Best Local Similarity 90.0%; Pred. NO. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
    |||||
Db 410 CGTGTCTGTGCTATGCC 391

```

RESULT 30

AG426940/c

LOCUS

```

DEFINITION
AG426940 Mus musculus molossinus DNA, clone:MSMG01-300H07.TJ, genomic survey
sequence.

```

AG426940
 AG426940.1 GI:48070003
 GSS.
 Mus musculus molossinus
 Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 BAC end Sequences of Library MSMg01
 2 (bases 1 to 972)
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 Direct Submission
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@riken.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the mouse BAC library MSMg01. For BAC
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp
 PRIMERS
 Sequencing : TJ
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.

FEATURES
 source
 1. .972
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clones="MSMg01-300H07.TJ"
 /sex="male"
 /tissue_type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 972;
 Best Local Similarity 90.0%; Pred. No. 1.9e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATCCC 20
 ||||| ||||| ||||| ||||| |||||
 Db 381 CGTGTGCTGCTAGTATCCC 362

RESULT 31
 BM048685/c
 LOCUS
 DEFINITION 603628283F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:5456637 5',
 mRNA sequence.

ACCESSION
 VERSION BM048685
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 994)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI955 row: e column: 22
 High quality sequence stop: 319.

FEATURES
 source

1. .994
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5456637"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 994;
 Best Local Similarity 90.0%; Pred. No. 1.9e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGCTAGTATCCC 20
 ||||| ||||| ||||| ||||| |||||
 Db 975 CGTGTGCTGCTAGTATCCC 956

RESULT 32
 CB994004/c

LOCUS
 DEFINITION AGENCOURT 13617214 NIH_MGC 148 Homo sapiens CDNA clone
 IMAGE:30333973 5', mRNA sequence.

ACCESSION
 VERSION CB994004.1 GI:30288524
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1010)
 NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Stefan Hanson
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM353 row: h column: 14
 High quality sequence stop: 246.

FEATURES
 source

1. .1010
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30333973"
 /tissue_type="pre-eclampsic placenta"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC 148"
 /note="Organ: placenta; Vector: pBluescriptR; Site_1:

all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 1010;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTCTAGTATCC 20
|||||
Db 617 CGTGTCTGTCTAGTATCC 598

RESULT 33

B0707308/c
LOCUS B0707308 1028 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_829192 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6280950
5', mRNA sequence.

ACCESSION B0707308
VERSION B0707308.1 GI:21846207

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1028)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L10M2472 row: p column: 07

High quality sequence stop: 570.

FEATURES

source

1..1028
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6280950"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 1028;
Best Local Similarity 90.0%; Pred. No. 1.9e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTCTAGTATCC 20
|||||
Db 642 CGTGTATCTTGTCTAGTATCC 623

RESULT 34

B1414634/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

B1414634 1263 bp mRNA linear EST 14-AUG-2001
602989764F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5145855 5',
mRNA sequence.

ACCESSION B1414634

VERSION B1414634.1 GI:15175557

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1263)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M11359 row: p column: 16

High quality sequence stop: 29

High quality sequence stop: 366.

Location/Qualifiers

1..1263

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:5145855"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NCI_CGAP_Lu33"

/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a

modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st

strand cDNA was prepared from mRNA obtained from pooled

lung tumors with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified p773 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 1263;

Best Local Similarity 90.0%; Pred. No. 1.9e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTCTAGTATCC 20
|||||
Db 858 CGTGTCTGTCTAGTATCC 839

RESULT 35

CF106699

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

Phytophthora infestans (potato late blight agent)

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phytophthora.

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

Phytophthora infestans

```

REFERENCE
AUTHORS      1 (bases 1 to 1560)
TITLE        Kin,K.S. and Judelson,H.S.
JOURNAL      Sporangia-specific gene expression in the oomycete phytopathogen
COMMENT      Phytophthora infestans
              Eukaryot. Cell 2 (6), 1376-1385 (2003)
              Contact: Judelson HS
              Department of Plant Pathology
              University of California
              Webber Hall, Riverside, CA 92521, USA
              Tel: 909 787 4199
              Fax: 909 787 4294
              Email: howard.judelson@ucr.edu
              mRNA induced in asexual sporangia compared to hyphae.

FEATURES
source
1..1560
/organism="Phytophthora infestans"
/mol_type="mRNA"
/strain="88069"
/db_xref="taxon:4787"
/dev_stage="Asexual sporangia from hyphae grown on rye agar"
/lab_host="DH10B"
/clone_libs="Undifferentiated asexual sporangia"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Assembly of sequence reads of overlapping cDNA clones"

ORIGIN
Query Match      84.0%; Score 16.8; DB 7; Length 1560;
Best Local Similarity 90.0%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 1479 CGTGTGCTGTGCTAGTCCC 1498

RESULT 36
AU076753/c
LOCUS          AU076753      178 bp      mRNA      linear      EST 04-MAY-2000
DEFINITION    AU076753 Sugano cDNA library Homo sapiens cDNA clone kaia0289
              similar to 5'-end region of Human CTLA4 counter-receptor (B7-2)
              mRNA, mRNA sequence.
ACCESSION     AU076753.1 GI:7439252
VERSION       AU076753.1
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H.,
              Tsunoda,T., Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A.
              and Sugano,S.
              Statistical analysis of the 5' untranslated region of human mRNA
              using 'Oligo-Capped' cDNA libraries
              Genomics 64 (3), 286-297 (2000)
              20221373
              PUBMED
              10756096
              Contact: Yutaka Suzuki
              Department of Virology
              Institute of Medical Science, University of Tokyo
              4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
              Email: ysuzuki@ims.u-tokyo.ac.jp
              Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
              Sugano,S. Construction and characterization of a full
              length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
              149-156 (1997)
              This clone was obtained from a 'full length-enriched' cDNA library
              constructed by 'oligo-capping' method. The coding region starts
              from the 50 bp upstream to the 3'-end.
              Location/Qualifiers
              1..178
              /organism="Homo sapiens"

FEATURES
source
1..178
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 47"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
              EcoRI; cDNA made by oligo-dT priming. Directionally
              cloned into EcoRI/XhoI sites using the following 5'
              adaptor: GGCACGAG(G). Size-selected >500bp for average
              insert size 1.8kb. Library constructed by Ling Hong in
              the laboratory of Gerald M. Rubin (University of
              California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and superscript II RT (Life Technologies).
              Note: this is a NIH_MGC Library."

```

```

ORIGIN
Query Match      82.0%; Score 16.4; DB 1; Length 178;
Best Local Similarity 85.0%; Pred. No. 2.5e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
    |||||
Db 79 CGGNGTCTGTGCTAGTGCC 60

RESULT 37
BM011282
LOCUS          BM011282      255 bp      mRNA      linear      EST 30-OCT-2001
DEFINITION    603635446F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5457923 5',
              mRNA sequence.
ACCESSION     BM011282.1 GI:16525636
VERSION       BM011282.1
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LHCMI958 row: k column: 12
              High quality sequence stop: 229.
              Location/Qualifiers
              1..255
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /tissue_type="neuroblastoma, cell line"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC 47"
              /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
              EcoRI; cDNA made by oligo-dT priming. Directionally
              cloned into EcoRI/XhoI sites using the following 5'
              adaptor: GGCACGAG(G). Size-selected >500bp for average
              insert size 1.8kb. Library constructed by Ling Hong in
              the laboratory of Gerald M. Rubin (University of
              California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and superscript II RT (Life Technologies).
              Note: this is a NIH_MGC Library."

ORIGIN
Query Match      82.0%; Score 16.4; DB 4; Length 255;
Best Local Similarity 89.5%; Pred. No. 2.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCCC 20
    |||||
Db 180 GTGTGCTGTGCTAGTCCC 198

RESULT 38
CR546210/c

```

LOCUS CR546210 411 bp mRNA linear EST 07-JUL-2004
DEFINITION DKFZp459A2113_r1_459 (synonym: pcor1) Pongo pygmaeus cDNA clone
ACCESSION DKFZp459A2113_5', mRNA sequence.
VERSION CR546210
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.
 1 (bases 1 to 411)
 Ansoerge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,
 Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
 Wiemann, S.
TITLE Pongo pygmaeus mRNA (Ansoerge, W., Krieger, S., Regiert, T., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert from S. Wiemann,
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);
 Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European
 Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp459A2113) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/projects/cdna/.
FEATURES Location/Qualifiers
 source
 1..411
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp459A2113"
 /tissue_type="cortex"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="459 (synonym: pcor1)"
 /note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"
ORIGIN
 Query Match 82.0%; Score 16.4; DB 7; Length 411;
 Best Local Similarity 94.4%; Pred. No. 2.7e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 TGTGCTCTGCTAGTCCC 20
 Db 177 TGTGCTCTGCTAGTCCC 160
 |||||
RESULT 39
EG408744/c 451 bp mRNA linear EST 13-MAR-2001
LOCUS gb80d06.y1 Moss EST library PPG Physcomitrella patens cDNA clone
DEFINITION PEP SOURCE ID:PPG_CopyA-100811_5', mRNA sequence.
ACCESSION EG408744
VERSION EG408744.1 GI:13315137
KEYWORDS EST.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
AUTHORS Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 451)
 Quatrano, R., Bashardes, S., Cove, D., Cuming, A., Knight, C.,
 Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T.,
 Underwood, K., Theising, B., Allen, M., Bowers, Y., Person, B.,
 Swaller, T., Steptoe, M., Gibbons, M., Harvey, N., Ritter, E.,
 Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
TITLE Leeds/Wash U Moss EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ralph Quatrano
 Leeds/Wash U Moss EST Project

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Libraries were constructed by Dr. Stavros Bashardes as part of the
 Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
 Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
 University Genome Sequencing Center For information on obtaining a
 clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 433.
 Location/Qualifiers
 source
 1..451
 /organism="Physcomitrella patens"
 /mol_type="mRNA"
 /db_xref="taxon:3218"
 /clone="PEP SOURCE ID:PPG_CopyA-100811"
 /tissue_type="gametophore: 30 day old tissue,
 ammonium-grown"
 /lab_host="DH10B"
 /clone_lib="Moss EST library PPG"
 /note="Vector: pAMP1; Construction of the cDNA library was
 performed by Dr. W. Gregg Clark using a modification of
 the cDNA synthesis protocol developed in the laboratory of
 Dr. Michael Lovett by Dr. Yulia Korshunova (personal
 communication). First polyA + RNA was isolated from total
 gametophore RNA using oligo dt magnetic beads. Following
 this, first strand cDNA synthesis was performed on the
 bead-bound polyA + RNA, during which an oligonucleotide
 anchor sequence was incorporated onto the 5'-ends of the
 cDNA. PCR amplification was then used to synthesize the
 second strand, to amplify the double stranded DNA, and to
 incorporate dUTP containing sequences into the ends of the
 double stranded cDNA. This DNA was size selected and
 cloned into pAMP1 using the CloneAMP pAMP1 System (Life
 Technologies, GibcoBRL) for cloning amplification products
 by a non-restriction site dependant process. The cloning
 was directional based on sequence asymmetry introduced at
 the ends during PCR amplification. The 3' cDNA ends are
 proximal to the NotI site of the multiple cloning site in
 pAMP1. This annealing mixture was transformed into
 chemically competent DH10B cells and selected for
 ampicillin resistant growth. The resulting clones (about
 330,000) were pooled to make the library."

ORIGIN
 Query Match 82.0%; Score 16.4; DB 4; Length 451;
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GTGTGCTCTGCTAGTCCC 19
 Db 155 GTGTGCTCTGCTAGTCCC 138
 |||||
RESULT 40
AW281714/c 481 bp mRNA linear EST 26-JUL-2002
LOCUS fj53h1.x1 zebrafish adult brain Danio rerio cDNA 3', mRNA
DEFINITION sequence.
ACCESSION AW281714
VERSION AW281714.1 GI:6670192
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 481)
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
 WashU Zebrafish EST Project 1998
 Unpublished (1998)
 Other ESTs: fj53h11.y1
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.edu
 cDNA Library Preparation: John Ngai. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone Distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)

Seq primer: -40UP
 High quality sequence stop: 440.

FEATURES

Location/Qualifiers
 1. .481
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /sex="mixed male and female"
 /tissue_type="brain"
 /dev_stage="adult"
 /lab_host="E. coli DH10B"
 /clone_lib="zebrafish adult brain"

/note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
 Original library was constructed in lambdaZIPLOX. Mass
 excision of the cDNA library was performed to yield
 pZIPLOX plasmids. Insert check was done in original
 library."

ORIGIN

Query Match 82.0%; Score 16.4; DB 2; Length 481;
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 TGTTGCTGCTAGTCC 20
 |||||
 Db 205 TGTTGCTGCTAGTCC 188

RESULT 41

BH111553/c
 LOCUS BH111553 497 bp DNA linear GSS 19-JUL-2001
 DEFINITION RPCI-24-252B2-TV RPCI-24 Mus musculus genomic clone RPCI-24-252B2,
 genomic survey sequence.

ACCESSION

VERSION BH111553

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 497)

REFERENCE

Authors Tsagave, G., Geer, K., Krol, M., Shvartsbeyn, A., Akinret, B., Levins, M.,
 Russell, D., de Jong, P. and Fraser, C.M.

TITLE

Mouse BAC End Sequences from Library RPCI-24

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC
 library availability, please contact Pieter de Jong
 (pjejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (http://www.choi.org/bacpac/orderingframe.html). BAC end
 page: http://ww.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 252 row: B column: 2
 Seq primer: T7
 Class: BAC ends

FEATURES

Location/Qualifiers
 1. .497
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-252B2"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /clone_lib="RPCI-24"
 /note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
 RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
 library was cloned in the pTARBAC1 cloning vector at the
 BamHI sites using MboI partially digested male C57BL/6J
 DNA."

ORIGIN

Query Match 82.0%; Score 16.4; DB 8; Length 497;
 Best Local Similarity 94.4%; Pred. No. 2.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGCTAGTCC 19

|||||

Db 246 GTGTGCTGCTAGTCC 229

RESULT 42

CD589669/c

LOCUS

CD589669 500 bp mRNA linear EST 16-JUN-2003
 DEFINITION RK053A3F03.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA
 clone RK053A3F03 5', mRNA sequence.

ACCESSION

VERSION CD589669

KEYWORDS CD589669.1 GI:31771021

SOURCE

EST.

ORGANISM

Danio rerio (zebrafish)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 500)

AUTHORS

Song, H.D., Wu, X.Y., Sun, X.J., Zhou, Y., Liu, T.X., Zhang, G.W.,
 Sheng, Y., Chen, Y., Ruan, Z., Jiang, C.L., Fan, H.Y., Look, A.T.,
 Zou, L.I. and Chen, Z.

TITLE

Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue

JOURNAL

Unpublished (2003)

COMMENT

Contact: Chen Z.

State Key Lab for Medical Genomics

Shanghai Institute of Hematology, Ruijin Hospital Affiliated to

Shanghai Second Medical University

197 Rui Jin Road II, Shanghai 200025, P. R. China

Tel: 86-21-64740490

Fax: 86-21-64743206

Email: zchen@stn.sh.cn

Seq primer: T3.

FEATURES

Location/Qualifiers

1. .500

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="RK053A3F03"

/dev_stage="mature"

/clone_lib="Zebrafish Kidney Marrow cDNA library"

/note="Organ: kidney; Vector: pBS-CMV; Site_1: XhoI;
 Site_2: EcoRI; Total RNA was extracted from the kidney

tissues of mature zebrafish. The poly (A)+ RNA fraction was separated from total RNA by oligo (dT) cellulose chromatography. Library was initially constructed in the lambdaZAP Express vector (Stratagene) and in vivo excised into pBS-CMV vector."

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 500;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTCTCTGTCTAGTCC 19
||||| ||||| ||||| ||||| |||||

Db 25 GTGTCTCTGTCTAGTCC 8

RESULT 43
AA067552 517 bp mRNA linear EST 06-NOV-1997
DEFINITION 26183 Lambda-PRL2 Arabidopsis thaliana cDNA clone 78C9T7, mRNA sequence.

ACCESSION AA067552
VERSION AA067552.1 GI:1565671
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 517)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.

TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
95148729
7846151

JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 2213tcn@bm.cl.msu.edu
Seq primer: 17 dye primer.

FEATURES

source
1..517
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="78C9T7"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site 1; Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dr primed cDNA."

ORIGIN

Query Match 82.0%; Score 16.4; DB 1; Length 517;
Best Local Similarity 89.5%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GTGTCTCTGTCTAGTCC 20
||||| ||||| ||||| ||||| |||||

Db 94 GTGTCTCTGTCTATTC 112

RESULT 44

CA041772/c 544 bp mRNA linear EST 04-MAR-2003
LOCUS ssalpinbs12124 gut Salmo salar cDNA, mRNA sequence.
DEFINITION CA041772
ACCESSION CA041772
VERSION CA041772.1 GI:24342692
KEYWORDS EST.
SOURCE Salmo salar (Atlantic salmon)
ORGANISM Salmo salar

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
1 (bases 1 to 544)
GRASP Consortium, Davidson,W.S., Koop,B.F. and
http://web.uvic.ca/cbr/grasp.
A survey of Salmo salar transcripts from high complexity cDNA libraries

TITLE Unpublished (2002)
JOURNAL Contact: Koop BF
COMMENT Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca

FEATURES

source
1..544
Location/Qualifiers
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone_lib="gut"
/note="Vector: pBluescriptII-SK+; Library Creator: Matthew L Rise ; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 544;
Best Local Similarity 94.4%; Pred. No. 2.8e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTCTCTGTCTAGTCC 19
||||| ||||| ||||| ||||| |||||

Db 284 GTGTCTCTGTCTAGTCC 267

RESULT 45

BM089797/c 578 bp mRNA linear EST 19-NOV-2001
LOCUS BM089797
DEFINITION 503647 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM089797
VERSION BM089797.1 GI:17000425
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 578)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,

TITLE
 Quackenbush, J. and Keele, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
JOURNAL
 Genome Res. 11 (4), 626-630 (2001)
MEDLINE
 21180013
PUBMED
 11282978
COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithem@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCCGTCAGTCAGCAG
 Plate: 5 row: D column: 18
 Seq primer: ATTGAGTGACACTATAG.

FEATURES
 source
 1..578
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 2BOV"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."

ORIGIN

Query Match 82.0%; Score 16.4; DB 4; Length 578;
 Best Local Similarity 94.4%; Pred.No. 2.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

Db 63 GTCTGTCTGTGCTAGTCC 46

RESULT 46
 CB504055/c
LOCUS
 CB504055.1 rev gut Salmo salar cDNA, mRNA sequence.
DEFINITION
 CB504055
ACCESSION
 CB504055.1 GI:29315281
VERSION
 EST.
KEYWORDS
 Salmo salar (Atlantic salmon)

SOURCE
 Salmo salar
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 1 (bases 1 to 592)

REFERENCE
 AUTHORS
 GRASP Consortium, Davidson, W.S., Koop, B.F. and
 http://web.uvic.ca/cbr/grasp.
TITLE
 A survey of Salmo salar transcripts from high complexity cDNA
 libraries

JOURNAL
 COMMENT
 Unpublished (2002)
 Contact: Koop BF
 Centre for Biomedical Research
 University of Victoria
 PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
 Tel: 250 472 4067
 Fax: 250 472 4075
 Email: bkoop@uvic.ca
 Genome Sciences Centre, BC Cancer Agency cDNA preparation,
 sequencing and bioinformatics: Y Butterfield, R Kirkpatrick, J
 Asano, N Girm, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prabhu,
 Smailius, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and
 M Marra.

FEATURES
 source
 1..592
 /organism="Salmo salar"
 /mol_type="mRNA"
 /strain="McConnell"
 /db_xref="taxon:8030"
 /clone_lib="gut"
 /note="Vector: pBlueScriptIIISK+; Library Creator: Matthew
 L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
 Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.); Simon
 Jones (PBS, Nanaimo, B.C.); Seaspring Hatchery (Crofton,
 B.C.), Rachel Roper (University of Victoria)"

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 592;
 Best Local Similarity 94.4%; Pred.No. 2.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

Db 395 GTGTGCTGTGCTAGTCC 378

RESULT 47
 CA041823
LOCUS
 CA041823.1 gut Salmo salar cDNA, mRNA sequence.
DEFINITION
 CA041823
ACCESSION
 CA041823.1 GI:24342743
VERSION
 EST.
KEYWORDS
 Salmo salar (Atlantic salmon)

SOURCE
 Salmo salar
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 1 (bases 1 to 630)

REFERENCE
 AUTHORS
 GRASP Consortium, Davidson, W.S., Koop, B.F. and
 http://web.uvic.ca/cbr/grasp.
TITLE
 A survey of Salmo salar transcripts from high complexity cDNA
 libraries

JOURNAL
 COMMENT
 Unpublished (2002)
 Contact: Koop BF
 Centre for Biomedical Research
 University of Victoria
 PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
 Tel: 250 472 4067
 Fax: 250 472 4075
 Email: bkoop@uvic.ca
 Genome Sciences Centre, BC Cancer Agency
 cDNA preparation, sequencing and bioinformatics:

Y Butterfield, R Kirkpatrick, J Asano, N Girm, R Guin, D Lee,
 S Lee, T Olson, P Pandoh, A Prabhu, D Smailius, L Spence, J Stott,
 S Taylor, G Yang, J Schein, S Jones and M Marra.
 POLYA=Yes.

FEATURES
 source
 1..630
 /organism="Salmo salar"
 /mol_type="mRNA"
 /strain="McConnell"
 /db_xref="taxon:8030"
 /clone_lib="gut"
 /note="Vector: pBlueScriptIIISK+; Library Creator: Matthew
 L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
 Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.); Simon
 Jones (PBS, Nanaimo, B.C.); Seaspring Hatchery (Crofton,
 B.C.), Rachel Roper (University of Victoria)"

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 630;
 Best Local Similarity 94.4%; Pred.No. 2.9e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTGCTGTGCTAGTCC 19

```

Db      517 GTGTCTCTGTCTGTCTCC 534
|||||
RESULT 48
CA061753
LOCUS      636 bp      mRNA      linear      EST 04-MAR-2003
DEFINITION      ssalrbs512251 mixed_tissue Salmo salar cDNA, mRNA sequence.
ACCESSION      CA061753
VERSION
KEYWORDS      EST.
SOURCE      Salmo salar (Atlantic salmon)
ORGANISM      Salmo salar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE      1 (bases 1 to 636)
AUTHORS      GRASP Consortium, Davidson,W.S., Koop,B.F. and
TITLE      http://web.uvic.ca/cbr/grasp.
A survey of Salmo salar transcripts from high complexity cDNA
libraries
JOURNAL      Unpublished (2002)
COMMENT      Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca
Genome Sciences Centre, BC Cancer Agency cDNA preparation,
sequencing and bioinformatics: Y Buterfield, R Kirkpatrick, J
Asano, N Girn, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prabhu, D
Smallus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and
M Marra.
POLYA=Yes.

FEATURES
source
1..636
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone_lib="gut"
/note="Vector: pBluescriptIISK+; Library Creator: Matthew
L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon
Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton,
B.C.), Rachel Roper (University of Victoria)"
ORIGIN
Query Match      82.0%; Score 16.4; DB 6; Length 648;
Best Local Similarity 94.4%; Pred. No. 2.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GTGTCTCTGTCTGTCTCC 19
|||||
Db      520 GTGTCTCTGTCTGTCTCC 537

RESULT 50
AL847928/C
LOCUS      657 bp      mRNA      linear      EST 26-NOV-2003
DEFINITION      AL847928 XGC-egg Xenopus tropicalis cDNA clone TEGG008d01 3', mRNA
sequence.
ACCESSION      AL847928
VERSION      AL847928.2 GI:38559484
KEYWORDS      EST.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM      Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE      1 (bases 1 to 657)
AUTHORS      Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE      Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL      Unpublished (2003)
COMMENT      On Sep 15, 2002 this sequence version replaced gi:22868193.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TEGG008d01.q1kT7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1..657
/organism="Xenopus tropicalis"

FEATURES
source
1..657
/organism="Xenopus tropicalis"

```

```
/mol_type="mrna"
/db_xref="taxon:8364"
/clone="TRgg008d01"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"
```

ORIGIN

```
Query Match      82.0%; Score 16.4; DB 1; Length 657;
Best Local Similarity 94.4%; Pred. No. 2.9e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      2 GTGTGTCGTGCTAGTCC 19
        |||||||
Db     217 GTGTGTCGTGCTAGTTC 200
```

Search completed: March 15, 2005, 21:29:05
Job time : 3146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 15:16:54 ; Search time 427 Seconds
(without alignments)
277.271 Million cell updates/sec

Title: US-09-980-953-256
Perfect score: 20
Sequence: 1 cgtgtgtctgtgtagtccg
Scoring table: IDENTITY_NUC
Gapop-10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

- Database : N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAJ33173	Aaf33173 Human B7-
2	20	100.0	20	ADJ27993	Ade27993 Human B7-
3	20	100.0	20	ADJ54435	Adj54435 Human B7-
C 4	20	100.0	124	AAJ01045	Aat01045 Human B7-
C 5	20	100.0	430	AAJ00427	Aac00427 Human sec
C 6	20	100.0	1120	AAJ81351	Aaq81351 Human B 1
C 7	20	100.0	1120	AAJ49181	Aat49181 Human B 1
C 8	20	100.0	1120	AAJ55784	Aav55784 Human B7-
C 9	20	100.0	1120	AAJ84049	Aac84049 Human B 1
C 10	20	100.0	1120	AAJ72340	Abv72340 Nucleotid
C 11	20	100.0	1120	AAJ27968	Aad27968 Human B7-
C 12	20	100.0	1120	AAJ60974	Aad60974 Human CD2
C 13	20	100.0	1120	ADK69859	Adk69859 Human B7-
C 14	20	100.0	1120	ADJ119363	Adj119363 Human B7-
C 15	20	100.0	2781	ADJ54476	Adj54476 Human B7-
C 16	20	100.0	2781	ADP10461	Adp10461 Reference
C 17	20	100.0	68001	ADJ54477	Adj54477 Human B7-
18	17.4	87.0	3168	ADN72456	Adn72456 Thale cre
19	16	80.0	368	AAJ05026	Aac05026 Human sec
20	16	80.0	372	AAJ01062	Aal01062 Human rep

21	16	80.0	372	4	ABL96529	AbL96529 Human tes
22	16	80.0	515	8	ABX91230	Abx91230 Murine ge
23	16	80.0	2175	6	ABZ21938	Abz21938 Human CAK
C 24	16	80.0	41434	11	ACN44256	Acn44256 Mouse gen
C 25	15.8	79.0	439	4	ABA09377	Aba09377 Human DNA
C 26	15.8	79.0	468	4	AAK82836	Aak82836 Human imm
C 27	15.8	79.0	1345	10	AAJ51678	Aaj51678 Human nuc
28	15.8	79.0	2304	6	ABZ11458	Abz11458 Human pol
29	15.8	79.0	2304	12	ADM43976	Adm43976 Novel hum
30	15.8	79.0	2434	4	AAI66465	Aai66465 P53 tumou
31	15.8	79.0	3545	4	ABL09670	AbL09670 Drosophil
32	15.8	79.0	5315	10	ADE55103	Ade55103 Rat gene
33	15.8	79.0	5856	4	ABL09674	AbL09674 Drosophil
34	15.8	79.0	82993	6	ABX09140	Abx09140 Mycobacte
35	15.8	79.0	110000	4	AAI99682	Aai99682 Continuat
36	15.8	79.0	110000	4	AAI99683	Aai99683 Continuat
37	15.8	79.0	214520	10	ADL13471	AdL13471 Osteoearth
38	15.8	79.0	347001	12	ADP43517	Adp43517 Human MAD
C 39	15.4	77.0	449	10	ADE59507	Ade59507 Rat gene
C 40	15.4	77.0	448	12	ADJ10940	Adj10940 Recombina
C 41	15.4	77.0	1644	6	ABZ12515	Abz12515 Arabidops
C 42	15.4	77.0	1690	6	ABL41879	AbL41879 Nucleotid
C 43	15.4	77.0	5819	2	AAJ20533	Aaj20533 Polynucle
C 44	15.4	77.0	21981	9	ADA02531	Ada02531 Mouse Sox
C 45	15.4	77.0	21981	10	ADB72269	AdB72269 Mouse Sox
C 46	15.4	77.0	21981	10	ADE95779	Ade95779 Mouse Sox
C 47	15.4	77.0	227448	13	ABD32841	Abd32841 Mouse can
48	15.4	77.0	229354	6	ABQ74179	AbQ74179 Human cmt
C 49	15.2	76.0	25	6	ACI35409	AcI35409 Human cmt
C 50	15.2	76.0	60	6	ABN39874	Abn39874 Human spl
51	15.2	76.0	327	6	ABL77238	AbL77238 Human ova
52	15.2	76.0	425	8	ABX44836	Abx44836 Bovine ES
C 53	15.2	76.0	461	9	ACH23811	Ach23811 Human adu
54	15.2	76.0	484	5	ADI77224	Adi77224 Human ova
55	15.2	76.0	484	5	ADI70911	Adi70911 Human ova
56	15.2	76.0	575	6	ABN73071	Abn73071 Bovine em
57	15.2	76.0	603	5	ADL42460	AdL42460 Human ova
58	15.2	76.0	647	5	AAJ67461	Aaj67461 DNA encod
C 59	15.2	76.0	1005	10	ADK68335	Adk68335 Human pri
C 60	15.2	76.0	1005	11	ADN39183	Adn39183 Cancer/an
C 61	15.2	76.0	1009	4	AAI59750	Aai59750 Human pol
C 62	15.2	76.0	1070	4	AAI57964	Aai57964 Human pol
C 63	15.2	76.0	1113	3	AAZ52422	Aaz52422 HTRM clon
64	15.2	76.0	1538	10	ADB53498	AdB53498 Primary r
65	15.2	76.0	1573	2	AAQ68222	Aaq68222 Neuroendo
C 66	15.2	76.0	1652	4	ABL15257	AbL15257 Drosophil
C 67	15.2	76.0	1740	4	ADJ08471	Adj08471 Human sec
C 68	15.2	76.0	1822	10	ADD44890	Add44890 Rat gene
C 69	15.2	76.0	1822	10	ADD44894	Add44894 Rat gene
C 70	15.2	76.0	3132	4	AAD08424	Aad08424 Human sec
C 71	15.2	76.0	3264	12	ADG76688	Adg76688 Phospholi
C 72	15.2	76.0	3287	4	ABL08582	AbL08582 Drosophil
C 73	15.2	76.0	3792	5	AAH73793	Aah73793 Human zin
C 74	15.2	76.0	3828	8	ABT18955	Abt18955 Aspergill
C 75	15.2	76.0	3853	6	ABQ54801	Abq54801 Human ova
C 76	15.2	76.0	4079	8	ABT18361	Abt18361 Aspergill
C 77	15.2	76.0	4172	4	ABL14128	AbL14128 Drosophil
C 78	15.2	76.0	4308	8	ABT20775	Abt20775 Aspergill
C 79	15.2	76.0	4311	4	ABL15256	AbL15256 Drosophil
C 80	15.2	76.0	4627	8	ADJ20177	Adj20177 Aspergill
C 81	15.2	76.0	5348	10	ADE62932	Ade62932 Human gen
C 82	15.2	76.0	6079	8	ABT17767	Abt17767 Aspergill
C 83	15.2	76.0	6089	12	ADJ95445	Adj95445 Mouse cDN
C 84	15.2	76.0	6089	13	ADS86867	AdS86867 cDNA enco
C 85	15.2	76.0	6108	4	ABL08478	AbL08478 Drosophil
C 86	15.2	76.0	6627	8	ABT19581	Abt19581 Aspergill
C 87	15.2	76.0	6808	4	ABL08488	AbL08488 Drosophil
C 88	15.2	76.0	16077	4	AAK86402	Aak86402 Human imm
C 89	15.2	76.0	17402	4	AAK68614	Aak68614 Human imm
C 90	15.2	76.0	17402	4	AAK86403	Aak86403 Human imm
C 91	15.2	76.0	24995	12	ADQ97269	Adq97269 Mouse can
92	15.2	76.0	30072	11	ACN44152	Acn44152 Mouse gen
C 93	15.2	76.0	50000	12	ADO44025	Ado44025 Nucleotid

94	15.2	76.0	73725	9	ADA02690	Human FLI	167	14.8	74.0	1873	4	AAS60756	Aas60756 Human can
95	15.2	76.0	73725	10	ADB72428	Human FLI	c 168	14.8	74.0	2999	4	ABL05137	Ab105137 Drosophil
96	15.2	76.0	73725	10	AD95938	Human FLI	c 169	14.8	74.0	3433	4	AAK81255	AAk81255 Human imm
97	15.2	76.0	80423	13	ABD32576	Mouse can	c 170	14.8	74.0	5647	4	ABL05136	Ab105136 Drosophil
98	15.2	76.0	110000	12	ADH69807_2	Continuation (3 of	171	14.8	74.0	7537	4	AAS36235	Aas36235 Human car
99	15.2	76.0	110000	12	ADH69807_3	Continuation (4 of	172	14.8	74.0	7537	4	AAK65643	AAk65643 Human imm
100	15.2	76.0	110000	13	ABD32923_1	Continuation (2 of	173	14.8	74.0	7537	10	AD646929	Ad646929 Human car
101	15.2	76.0	120144	13	ABD33161_1	Abd33161 Murine can	174	14.8	74.0	7537	13	ADJ08347	Adj08347 Human car
102	15.2	76.0	143040	11	ABD20648	Human pul	c 175	14.8	74.0	8047	4	AAK81256	AAk81256 Human imm
103	15.2	76.0	143067	11	ABD20815	Abd20815 Human pul	c 176	14.8	74.0	8047	4	AAK81256	AAk81256 Human imm
104	15.2	76.0	143068	11	ABD34983	Abd34983 Human pul	c 177	14.8	74.0	8047	4	AAK81256	AAk81256 Human imm
105	15.2	76.0	143068	3	AAA35150	Human ade	c 178	14.8	74.0	13066	4	ABL12848	Ab112848 Drosophil
106	15.2	76.0	143068	3	AAA35150	Human ade	c 179	14.8	74.0	29411	4	AAK76613	AAk76613 Human imm
107	15.2	76.0	143068	3	AAf211272	Human low	179	14.8	74.0	40304	9	ADA03014	Ada03014 Human NCF
108	15.2	76.0	143068	3	AAf211105	Human low	180	14.8	74.0	40304	10	ADB272752	Adb272752 Human NCF
109	15.2	76.0	143068	6	ABL68124	Ovary can	181	14.8	74.0	40304	10	ADC85494	Adc85494 Human Ncf
110	15.2	76.0	143068	6	ABT11034	Abt11034 Human bre	182	14.8	74.0	40304	12	ADM74609	Adm74609 Human car
111	15.2	76.0	143068	10	ABZ96966	Human nuc	c 183	14.8	74.0	40429	10	ADB74376	Adb74376 Mycobacte
112	15.2	76.0	143068	10	ABZ96966	Human nuc	184	14.8	74.0	41322	9	AAAL62633	Aal62633 Human CD3
113	15.2	76.0	149412	3	AAA35151	Human ade	185	14.8	74.0	44456	4	ABL09962	Ab109962 Drosophil
114	15.2	76.0	152739	11	ABD20816	Abd20816 Human pul	186	14.8	74.0	44990	9	ADA02711	Ada02711 Mouse Vda
115	15.2	76.0	152740	3	AAf211273	Human low	187	14.8	74.0	44990	10	ADB72449	Adb72449 Mouse Vda
116	15.2	76.0	152740	3	AAf211273	Human low	188	14.8	74.0	44990	10	ADB72449	Adb72449 Mouse Vda
117	15	75.0	291	3	AAc29169	Abc29169 Human nuc	189	14.8	74.0	48436	6	ABN89533	Abn89533 Human cor
118	15	75.0	532	1	AAm50221	BamHI-Eco	c 190	14.8	74.0	60935	13	ACN37224	Acn37224 Human per
119	15	75.0	533	1	AAm60422	Restricti	191	14.8	74.0	86574	6	ABK83560	Abk83560 Human cDN
120	15	75.0	533	1	AAm60975	PHOSR pro	192	14.8	74.0	86574	13	ADRS2822	Adrs2822 Drug ther
121	15	75.0	539	1	AAm50222	BamHI-Eco	193	14.8	74.0	87869	12	ADP13458	Adp13458 Renal cel
122	15	75.0	539	1	AAm60423	Restricti	194	14.8	74.0	106938	13	ABD33432	Abd33432 Human can
123	15	75.0	539	1	AAm60976	PHOSY pro	195	14.8	74.0	106938	13	ABD33432	Abd33432 Human can
124	15	75.0	541	13	ADR49636	PHOS pro	196	14.8	74.0	110000	6	ABK83560	Abk83560 Human cDN
125	15	75.0	546	6	ABN71067	Streptoco	197	14.8	74.0	110000	12	ADN97989_05	Adn97989_05 Continuation (6 of
126	15	75.0	558	6	ABN67141	Streptoco	198	14.8	74.0	110000	12	ADN97989_05	Adn97989_05 Continuation (6 of
127	15	75.0	610	1	AAm70940	DNA sequ	199	14.8	74.0	110000	12	ADN97989_05	Adn97989_05 Continuation (6 of
128	15	75.0	612	2	AAQ52326	Sequence	200	14.8	74.0	110000	13	ABD32627_0	Abd32627 Mouse can
129	15	75.0	622	1	AAm40105	Sequence	201	14.8	74.0	110000	13	ABD32627_0	Abd32627 Mouse can
130	15	75.0	622	1	AAm50217	DNA sequ	202	14.8	74.0	160552	4	AAAD02697	Aad02697 Human gly
131	15	75.0	627	1	AAm70086	Sequence	203	14.8	74.0	175338	11	ACN45088	Acn45088 Mouse gen
132	15	75.0	629	1	AAm70320	Sequence	204	14.8	74.0	188888	6	ABQ75562	Abq75562 Human rel
133	15	75.0	721	10	ADD16755	ADD16755 DNA (SeqI	c 205	14.8	74.0	193303	12	ADF13112	Adf13112 Hypermeth
134	15	75.0	721	10	ADD17580	ADD17580 DNA (SeqI	c 206	14.8	74.0	193303	12	ADF13112	Adf13112 Hypermeth
135	15	75.0	1566	2	AAQ3768	Alpha-int	c 207	14.8	74.0	193303	12	ADI37268	Adi37268 Hypermeth
136	15	75.0	1904	4	AAH29723	S cerevis	c 208	14.8	74.0	193303	12	ADI37261	Adi37261 Hypermeth
137	15	75.0	8140	2	AAV23013	Nucleic a	c 209	14.8	74.0	197997	10	AAAL54074	Aal54074 Human tra
138	15	75.0	110000	6	ABN71527_15	Continuation (16 o	c 210	14.8	74.0	219352	13	ABD33098	Abd33098 Murine ca
139	15	75.0	121124	12	ABQ97107	Mouse can	211	14.8	74.0	277616	13	ABD32602	Abd32602 Human can
140	14.8	74.0	216	6	ABQ92911	Triticum	c 212	14.4	72.0	271	12	ACH81276	Ach81276 Human gen
141	14.8	74.0	232	2	AAQ77572	Human gen	c 213	14.4	72.0	296	12	ADQ21051	Adq21051 Human sof
142	14.8	74.0	394	6	ABN66018	Human can	c 214	14.4	72.0	315	5	ABV06855	Abv06855 Human pro
143	14.8	74.0	408	9	ADA29682	DNA encod	c 215	14.4	72.0	328	5	ABV06556	Abv06556 Human pro
144	14.8	74.0	519	9	ACH18732	Human adu	c 216	14.4	72.0	404	12	ADQ18454	Adq18454 Human sof
145	14.8	74.0	529	4	AAI11760	Probe #70	c 217	14.4	72.0	416	8	ABX62349	Abx62349 Arabidops
146	14.8	74.0	529	4	ABAA61665	Human foe	c 218	14.4	72.0	420	5	ABV36507	Abv36507 Human pro
147	14.8	74.0	529	4	AAI41577	Probe #10	c 219	14.4	72.0	445	8	ABX43581	Abx43581 Bovine ES
148	14.8	74.0	529	4	ABA29321	Probe #77	c 220	14.4	72.0	470	9	ACH41379	Ach41379 Human foe
149	14.8	74.0	529	4	AAK35861	Human bon	c 221	14.4	72.0	472	5	ABV36800	Abv36800 Human pro
150	14.8	74.0	529	4	AAK09967	Human bra	c 222	14.4	72.0	490	9	ACH13691	Ach13691 Human adu
151	14.8	74.0	529	4	AB355570	Human liv	c 223	14.4	72.0	520	5	ABV06555	Abv06555 Human pro
152	14.8	74.0	554	2	AAAD21044	Polyucle	c 224	14.4	72.0	525	12	ACH96688	Ach96688 Human gen
153	14.8	74.0	601	10	AAD63646	Soybean d	c 225	14.4	72.0	547	12	ACH67573	Ach67573 Human gen
154	14.8	74.0	601	10	ADH63018	Soybean d	c 226	14.4	72.0	585	3	AAAC49903	Aac49903 Arabidops
155	14.8	74.0	620	12	ADQ18837	Human sof	c 227	14.4	72.0	611	10	ADE62612	Ad62612 Rat gene
156	14.8	74.0	677	10	ADD34870	Mouse mit	c 228	14.4	72.0	676	5	ABA13949	Ab13949 Human ner
157	14.8	74.0	705	10	AAAD63637	Soybean d	c 229	14.4	72.0	749	5	ABL11761	Ab11761 Drosophil
158	14.8	74.0	705	10	AAAD63637	Soybean d	c 230	14.4	72.0	760	3	AAA14071	Aaa14071 Human SPR
159	14.8	74.0	705	10	ADH63000	Soybean d	c 231	14.4	72.0	882	5	ABV21146	Abv21146 Human pro
160	14.8	74.0	815	6	AB199287	Mouse isc	c 232	14.4	72.0	882	5	ABV21146	Abv21146 Human pro
161	14.8	74.0	855	3	AA250926	Human pro	c 233	14.4	72.0	985	3	AAAC49911	Aac49911 Arabidops
162	14.8	74.0	955	12	ADP02970	Human hou	c 234	14.4	72.0	990	3	AAAC37904	Aac37904 Arabidops
163	14.8	74.0	955	13	AD588468	Human hou	c 235	14.4	72.0	1071	3	AAAC49910	Aac49910 Arabidops
164	14.8	74.0	966	4	AAI88604	Human pol	c 236	14.4	72.0	1076	3	AAAC40666	Aac40666 Arabidops
165	14.8	74.0	1311	13	ADT48577	Bacterial	c 237	14.4	72.0	1178	3	AAZ3696	Aaz3696 Zea maye
166	14.8	74.0	1873	4	AAS60755	Human can	c 238	14.4	72.0	1237	3	AAZ98072	Aaz98072 Human sec

240	14.4	72.0	1237	4	AD11685	Human sec	Ad11685 Human sec	c 313	14.4	72.0	25346	12	ADQ74844	AdQ74844 Loline al
241	14.4	72.0	1237	6	ABK69781	Human sec	Adk69781 Human sec	314	14.4	72.0	25806	4	ABK67666	Abk67666 Human imm
242	14.4	72.0	1237	8	ADA40473	Human sec	Ada40473 Human sec	315	14.4	72.0	29376	4	ABL08834	AbL08834 Drosophil
243	14.4	72.0	1237	8	ACC50815	Human sec	Acc50815 Human sec	c 316	14.4	72.0	32134	5	AAI199172	Aai199172 Human exc
244	14.4	72.0	1237	8	ABZ71451	Secreted	Abz71451 Secreted	c 317	14.4	72.0	32134	5	AAI163522	Aai163522 Human kid
245	14.4	72.0	1237	9	ADB91410	Human sec	Adb91410 Human sec	c 318	14.4	72.0	32192	4	AAI199173	Aai199173 Human exc
246	14.4	72.0	1237	10	ADC73392	Human sec	Adc73392 Human sec	c 319	14.4	72.0	32192	5	AAI163523	Aai163523 Human kid
247	14.4	72.0	1237	10	ADA56633	Gene enco	Ada56633 Gene enco	c 320	14.4	72.0	32763	4	AAK68779	Aak68779 Human imm
248	14.4	72.0	1251	8	ADA40097	Human sec	Ada40097 Human sec	c 321	14.4	72.0	35100	4	AAK69767	Aak69767 Human imm
249	14.4	72.0	1251	8	ACC50601	Human sec	Acc50601 Human sec	c 322	14.4	72.0	35100	4	AAK65700	Aak65700 Human imm
250	14.4	72.0	1251	8	ABZ71329	Secreted	Abz71329 Secreted	c 323	14.4	72.0	35115	4	AAK69766	Aak69766 Human imm
251	14.4	72.0	1251	9	ADB91255	Human sec	Adb91255 Human sec	c 324	14.4	72.0	35115	4	AAK65699	Aak65699 Human imm
252	14.4	72.0	1251	10	ADC73672	Human sec	Adc73672 Human sec	325	14.4	72.0	35460	12	ADQ59458	AdQ59458 Human can
253	14.4	72.0	1251	10	ADA56271	Gene enco	Ada56271 Gene enco	c 326	14.4	72.0	36785	4	AAK82208	Aak82208 Human imm
254	14.4	72.0	1340	5	ABV25507	Human pro	Abv25507 Human pro	327	14.4	72.0	70549	12	ADQ97436	AdQ97436 Mouse can
255	14.4	72.0	1340	5	ABV24187	Human pro	Abv24187 Human pro	c 328	14.4	72.0	73930	12	ADQ97577	AdQ97577 Human can
256	14.4	72.0	1412	5	ABV24402	Human pro	Abv24402 Human pro	c 329	14.4	72.0	79731	12	ADQ97640	AdQ97640 Mouse can
257	14.4	72.0	1412	5	ABV24497	Human pro	Abv24497 Human pro	330	14.4	72.0	95400	12	ADP08388	Adp08388 Human lam
258	14.4	72.0	1639	12	ADU12794	Human pro	Adu12794 Human pro	c 331	14.4	72.0	96276	12	ADQ17634	AdQ17634 Human sof
259	14.4	72.0	1681	4	AAH17412	Human cdn	Aah17412 Human cdn	332	14.4	72.0	96596	9	ADA02504	Ada02504 Human BAC
260	14.4	72.0	1698	6	ABK69901	Human sec	Abk69901 Human sec	333	14.4	72.0	96596	10	ADB72242	AdB72242 Human BAC
261	14.4	72.0	1740	11	ADM03148	Human cdn	Adm03148 Human cdn	334	14.4	72.0	96596	10	ABE95752	ABe95752 Human BAC
262	14.4	72.0	2030	3	AAAL4074	Human SPR	Aaal4074 Human SPR	335	14.4	72.0	110000	10	ABQ84281_2	ABq84281_2 of
263	14.4	72.0	2043	6	ABK74362	Bacillus	Abk74362 Bacillus	336	14.4	72.0	110000	10	ABQ84281_3	ABq84281_3 of
264	14.4	72.0	2138	3	AAAL4073	Human SPR	Aaal4073 Human SPR	c 337	14.4	72.0	168198	12	ADQ59452	AdQ59452 Human can
265	14.4	72.0	2195	3	AAAL4072	Human SPR	Aaal4072 Human SPR	c 338	14.4	72.0	173810	6	ABN85752	ABn85752 Mouse chr
266	14.4	72.0	2199	4	AAI59630	Human pol	Aai59630 Human pol	c 339	14.4	72.0	176080	12	ADL08124	AdL08124 Human gen
267	14.4	72.0	2211	5	ABAL1513	Human ner	Abal1513 Human ner	340	14.4	72.0	201143	6	ABK83568	ABk83568 Human DNA
268	14.4	72.0	2254	3	ABCT8059	Human can	Abct8059 Human can	341	14.4	72.0	208765	12	ADQ97430	AdQ97430 Mouse can
269	14.4	72.0	2330	4	AAK32646	Human can	Aak32646 Human can	342	14.4	72.0	247544	12	ADQ59419	AdQ59419 Human can
270	14.4	72.0	2390	6	ABQ76001	Human ger	Abq76001 Human ger	c 343	14.4	72.0	264965	12	ADN16203	AdN16203 Human sul
271	14.4	72.0	2393	4	AAI57844	Human pol	Aai57844 Human pol	c 344	14.4	72.0	268685	6	ABS56563	ABs56563 Human SUL
272	14.4	72.0	2483	10	ADB62098	Human cdn	Adb62098 Human cdn	345	14.4	72.0	310122	13	ABD32533	ABd32533 Mouse can
273	14.4	72.0	2485	5	ABAL17501	Human ner	Abal17501 Human ner	346	14.2	71.0	25	9	ACI07954	ACi07954 Human mic
274	14.4	72.0	2485	5	ABAL15472	Human ner	Abal15472 Human ner	c 347	14.2	71.0	33	3	AAZ96553	Aaz96553 T cell an
275	14.4	72.0	2485	5	ABAL17503	Human ner	Abal17503 Human ner	348	14.2	71.0	50	6	ABZ00190	ABz00190 Human leu
276	14.4	72.0	2485	5	ABAL16616	Human ner	Abal16616 Human ner	c 349	14.2	71.0	54	2	AAK67219	Aak67219 Mouse CD4
277	14.4	72.0	2485	5	ABAL16617	Human ner	Abal16617 Human ner	350	14.2	71.0	95	4	AAI24619	Aai24619 Probe #14
278	14.4	72.0	2530	12	ADQ22948	Human sof	Adq22948 Human sof	351	14.2	71.0	95	4	ABA69823	ABa69823 Human foe
279	14.4	72.0	2555	3	ACAT75393	Human ORF	Acat75393 Human ORF	352	14.2	71.0	95	4	AAI49921	Aai49921 Probe #18
280	14.4	72.0	2556	4	ABL20450	Drosophil	AbL20450 Drosophil	353	14.2	71.0	95	4	ABA36691	ABa36691 Probe #15
281	14.4	72.0	2558	10	ABE99949	Bacterial	ABe99949 Bacterial	354	14.2	71.0	95	4	AAK43911	Aak43911 Human bon
282	14.4	72.0	2717	5	AAH48361	Aldehyde-	Aah48361 Aldehyde-	355	14.2	71.0	95	4	AAK43911	Aak43911 Human bon
283	14.4	72.0	2725	6	ABL67209	Thyroid c	AbL67209 Thyroid c	356	14.2	71.0	95	4	AAK18025	Aak18025 Human bra
284	14.4	72.0	2725	6	ABN95683	Gene #218	ABn95683 Gene #218	357	14.2	71.0	95	4	ABS43564	ABs43564 Human liv
285	14.4	72.0	2834	6	ABA05986	Human bra	AbA05986 Human bra	358	14.2	71.0	95	6	ABS18141	ABs18141 Human gen
286	14.4	72.0	2903	4	ABL17160	Drosophil	AbL17160 Drosophil	c 358	14.2	71.0	121	10	ADH92823	ADh92823 Human gen
287	14.4	72.0	3138	10	ADC32235	Human nov	AdC32235 Human nov	359	14.2	71.0	133	3	AAQ4914	AAq4914 Human sec
288	14.4	72.0	3182	2	AAQ70730	TATA-bind	Aaq70730 TATA-bind	360	14.2	71.0	148	11	ADM13155	AdM13155 Human ees
289	14.4	72.0	3182	2	AAAT42216	Human TAR	Aat42216 Human TAR	c 361	14.2	71.0	159	3	AAZ42384	Aaz42384 Human 5'
290	14.4	72.0	3182	2	AAAT79593	TATA-bind	Aat79593 TATA-bind	c 362	14.2	71.0	161	2	AAQ14558	Aaq14558 HCMV AD16
291	14.4	72.0	3202	12	ADQ22654	Human sof	AdQ22654 Human sof	c 363	14.2	71.0	177	8	ACD05668	ACd05668 cDNA enco
292	14.4	72.0	3202	12	ADQ22654	Breast ca	AdQ22654 Breast ca	c 364	14.2	71.0	190	2	AAH85712	AAh85712 Human ein
293	14.4	72.0	3202	12	ADQ22654	Human nuc	AdQ22654 Human nuc	365	14.2	71.0	216	9	ACH45516	ACH45516 Human foe
294	14.4	72.0	3202	12	ADQ22654	Human nuc	AdQ22654 Human nuc	366	14.2	71.0	219	3	AACT1382	AAcT1382 Human sec
295	14.4	72.0	3202	12	ADQ22654	Human nuc	AdQ22654 Human nuc	c 367	14.2	71.0	223	4	AAI20560	Aai20560 Probe #10
296	14.4	72.0	3202	12	ADQ22654	Toxicity	AdQ22654 Toxicity	368	14.2	71.0	223	4	ABA65603	ABa65603 Human foe
297	14.4	72.0	3202	12	ADQ22654	Human cel	AdQ22654 Human cel	369	14.2	71.0	223	4	AAI45770	Aai45770 Probe #14
298	14.4	72.0	3202	12	ADQ22654	Gene #165	AdQ22654 Gene #165	370	14.2	71.0	223	4	ABA47709	ABa47709 Human bre
299	14.4	72.0	3202	12	ADQ22654	Human tum	AdQ22654 Human tum	371	14.2	71.0	223	4	ABA32691	ABa32691 Probe #11
300	14.4	72.0	3202	12	ADQ22654	Human sof	AdQ22654 Human sof	372	14.2	71.0	223	4	AAK39754	Aak39754 Human bon
301	14.4	72.0	3202	12	ADQ22654	Human pot	AdQ22654 Human pot	373	14.2	71.0	223	4	AAK14006	Aak14006 Human bra
302	14.4	72.0	3202	12	ADQ22654	Human rep	AdQ22654 Human rep	374	14.2	71.0	223	4	ABS39341	ABs39341 Human liv
303	14.4	72.0	3202	12	ADQ22654	Drosophil	AbL28998 Drosophil	375	14.2	71.0	223	5	AAI06259	AAi06259 Probe #62
304	14.4	72.0	3202	12	ADQ22654	Human rep	AbL28998 Human rep	376	14.2	71.0	223	3	ABS13848	ABs13848 Human gen
305	14.4	72.0	3202	12	ADQ22654	Drosophil	AbL28998 Drosophil	c 377	14.2	71.0	238	3	AACT1351	AAcT1351 Human sec
306	14.4	72.0	3202	12	ADQ22654	Drosophil	AbL28998 Drosophil	c 378	14.2	71.0	240	6	ABN16239	ABn16239 Human ORF
307	14.4	72.0	3202	12	ADQ22654	Human CD3	Aaf63406 Human CD3	c 379	14.2	71.0	245	8	AAI29027	Aai29027 Colon tum
308	14.4	72.0	3202	12	ADQ22654	Human CD3	Aaf63407 Human CD3	c 380	14.2	71.0	245	8	ABZ33213	ABz33213 Human col
309	14.4	72.0	3202	12	ADQ22654	Human int	Abk72689 Human int	c 381	14.2	71.0	264	2	AAT23660	Aat23660 Human gen
310	14.4	72.0	3202	12	ADQ22654	Human int	Abk15357 Human int	c 382	14.2	71.0	289	6	ABN17501	ABn17501 Human ORF
311	14.4	72.0	3202	12	ADQ22654	Mammalian	AdL19368 Mammalian	c 383	14.2	71.0	294	6	ABN19677	ABn19677 Human ORF
312	14.4	72.0	3202	12	ADQ22654	Human imm	Aak86765 Human imm	c 384	14.2	71.0	303	4	AAI82002	Aai82002 Human pol
313	14.4	72.0	3202	12	ADQ22654	Human GPC	AdC86168 Human GPC	c 385	14.2	71.0	314	9	ACH31804	ACH31804 Human bon

C 386 14.2 71.0 319 4 AAL13361 Aal13361 Human bre
C 387 14.2 71.0 320 3 AAZ38043 Aaz38043 T Lymphoc
C 388 14.2 71.0 326 4 AAI18844 Aai18844 Human pol
C 389 14.2 71.0 332 4 AAI80242 Aai80242 Human pol
C 390 14.2 71.0 335 4 AAI23894 Aai23894 Probe #13
C 391 14.2 71.0 335 4 ABA69012 Aab69012 Human foe
C 392 14.2 71.0 335 4 AAI49202 Aai49202 Probe #17
C 393 14.2 71.0 335 4 ABA51022 Aab51022 Human bre
C 394 14.2 71.0 335 4 ABA35957 Aab35957 Probe #14
C 395 14.2 71.0 335 4 AAK43125 Aak43125 Human bon
C 396 14.2 71.0 335 4 AAK17332 Aak17332 Human bra
C 397 14.2 71.0 335 4 ABA2757 Aab2757 Human liv
C 398 14.2 71.0 335 5 AAI09495 Aai09495 Probe #94
C 399 14.2 71.0 335 6 ABA17200 Aab17200 Human gen
C 400 14.2 71.0 345 8 ABA21756 Aab21756 S2 subtra
C 401 14.2 71.0 350 4 AAL11801 Aal11801 Human bre
C 402 14.2 71.0 350 4 AAI18844 Aai18844 Human pol
C 403 14.2 71.0 351 4 AAL20691 Aal20691 Human bre
C 404 14.2 71.0 358 4 AAI14700 Aai14700 Probe #46
C 405 14.2 71.0 358 4 ABA56431 Aab56431 Human foe
C 406 14.2 71.0 358 4 AAI36067 Aai36067 Probe #47
C 407 14.2 71.0 358 4 ABA45904 Aab45904 Human bre
C 408 14.2 71.0 358 4 ABA26064 Aab26064 Probe #45
C 409 14.2 71.0 358 4 AAK30104 Aak30104 Human bon
C 410 14.2 71.0 358 4 AAK04597 Aak04597 Human bra
C 411 14.2 71.0 358 4 ABA29752 Aab29752 Human liv
C 412 14.2 71.0 358 5 AAI04505 Aai04505 Probe #44
C 413 14.2 71.0 358 6 ABA20305 Aab20305 Human ORF
C 414 14.2 71.0 358 6 ABA504684 Aab504684 Human gen
C 415 14.2 71.0 361 4 AAI10753 Aai10753 Human bre
C 416 14.2 71.0 362 4 AAI17828 Aai17828 Human bre
C 417 14.2 71.0 362 8 ABA280524 Aab280524 Novel exp
C 418 14.2 71.0 394 6 ABAQ92737 Aabq92737 Triticum
C 419 14.2 71.0 403 11 ABA181999 Aab181999 Breast ca
C 420 14.2 71.0 408 6 ABL82135 Aab182135 Human ova
C 421 14.2 71.0 412 12 ADL18342 Aab18342 Killifish
C 422 14.2 71.0 417 12 ADO19926 Aab19926 Human PRO
C 423 14.2 71.0 423 4 AAI15387 Aai15387 Probe #53
C 424 14.2 71.0 423 4 ABA57189 Aab57189 Human foe
C 425 14.2 71.0 423 4 AAI36719 Aai36719 Probe #54
C 426 14.2 71.0 423 4 ABA26767 Aab26767 Probe #52
C 427 14.2 71.0 423 4 AAK30806 Aak30806 Human bon
C 428 14.2 71.0 423 4 AAK05217 Aak05217 Human bra
C 429 14.2 71.0 423 4 ABA30483 Aab30483 Human liv
C 430 14.2 71.0 423 6 ABA505551 Aab505551 Human gen
C 431 14.2 71.0 432 3 AAK04197 Aak04197 Human sec
C 432 14.2 71.0 436 6 ABL65362 Aab165362 Lung canc
C 433 14.2 71.0 436 6 ABL64723 Aab164723 Lung canc
C 434 14.2 71.0 441 5 ABAV16075 Aabv16075 Human pro
C 435 14.2 71.0 442 13 ABA099950 Aab099950 Human the
C 436 14.2 71.0 445 12 ADO63314 Aab063314 Transcript
C 437 14.2 71.0 449 4 AAK59387 Aak59387 Human imm
C 438 14.2 71.0 456 10 ABA284709 Aab284709 Toxicolog
C 439 14.2 71.0 463 9 ABA34826 Aab34826 Human end
C 440 14.2 71.0 467 4 AAI11345 Aai11345 Probe #12
C 441 14.2 71.0 467 4 ABA53009 Aab53009 Human foe
C 442 14.2 71.0 467 4 AAI32614 Aai32614 Probe #13
C 443 14.2 71.0 467 4 ABA42580 Aab42580 Human bre
C 444 14.2 71.0 467 4 ABA22784 Aab22784 Probe #12
C 445 14.2 71.0 467 4 AAK26719 Aak26719 Human bon
C 446 14.2 71.0 467 4 AAK01257 Aak01257 Human bra
C 447 14.2 71.0 467 4 ABA26308 Aab26308 Human liv
C 448 14.2 71.0 467 5 AAI01261 Aai01261 Probe #12
C 449 14.2 71.0 467 6 ABA01313 Aab01313 Human gen
C 450 14.2 71.0 470 4 AAI16476 Aai16476 Probe #64
C 451 14.2 71.0 470 4 ABA59454 Aab59454 Human foe
C 452 14.2 71.0 470 4 AAI39277 Aai39277 Probe #79
C 453 14.2 71.0 470 4 ABA28111 Aab28111 Probe #65
C 454 14.2 71.0 470 4 AAK33508 Aak33508 Human bon
C 455 14.2 71.0 470 4 AAK07692 Aak07692 Human bra
C 456 14.2 71.0 470 4 ABA33268 Aab33268 Human liv
C 457 14.2 71.0 470 6 ABA08354 Aab08354 Human gen
C 458 14.2 71.0 471 8 ABA38704 Aab38704 Bovine ES

C 459 14.2 71.0 473 6 ABAQ57152 Aabq57152 Human col
C 460 14.2 71.0 478 9 ACHI15546 Achi15546 Human adu
C 461 14.2 71.0 486 11 ABD16922 Aab16922 Pseudomon
C 462 14.2 71.0 495 9 ACH50676 Aach50676 Human mam
C 463 14.2 71.0 499 9 ACH34345 Aach34345 Human end
C 464 14.2 71.0 504 4 AAI17874 Aai17874 Human bre
C 465 14.2 71.0 524 5 ABAV15836 Aabv15836 Human pro
C 466 14.2 71.0 541 4 AHA29315 Aah29315 Drosophil
C 467 14.2 71.0 553 10 ADP19831 Aadd19831 Seabass p
C 468 14.2 71.0 556 5 ABAV30326 Aabv30326 Human pro
C 469 14.2 71.0 556 5 ABAV30142 Aabv30142 Human pro
C 470 14.2 71.0 560 1 AAN60553 Aaan60553 Bovine ma
C 471 14.2 71.0 560 2 AAO6235 Aaao6235 Fragment
C 472 14.2 71.0 575 5 ABAV50055 Aabv50055 Human pro
C 473 14.2 71.0 576 2 AAO83877 Aaao83877 Hepatitis
C 474 14.2 71.0 576 2 AAO83876 Aaao83876 Hepatitis
C 475 14.2 71.0 576 2 AAO83875 Aaao83875 Hepatitis
C 476 14.2 71.0 576 2 AAT16589 Aaat16589 Hepatitis
C 477 14.2 71.0 576 2 AAT16590 Aaat16590 Hepatitis
C 478 14.2 71.0 576 2 AAT16591 Aaat16591 Hepatitis
C 479 14.2 71.0 577 5 ABAV50123 Aabv50123 Human pro
C 480 14.2 71.0 591 8 AAD48979 Aaad48979 Human met
C 481 14.2 71.0 592 11 ACR86769 Aacr86769 Breast ca
C 482 14.2 71.0 601 11 ACR86828 Aacr86828 Breast ca
C 483 14.2 71.0 605 4 AAH06164 Aah06164 Human CDN
C 484 14.2 71.0 611 10 ADC74843 Aadc74843 Human pro
C 485 14.2 71.0 618 11 ACR87863 Aacr87863 Breast ca
C 486 14.2 71.0 624 13 ADT45043 Aadt45043 Bacterial
C 487 14.2 71.0 628 5 ABAI2571 Aabai2571 Human ner
C 488 14.2 71.0 633 12 ADJ86984 Adj86984 Synthetic
C 489 14.2 71.0 637 10 ADK54387 Adk54387 Plant DNA
C 490 14.2 71.0 643 11 ACR80950 Aacr80950 Breast ca
C 491 14.2 71.0 709 4 AAL22231 Aaal22231 Human bre
C 492 14.2 71.0 713 5 ABAV45636 Aabv45636 Human pro
C 493 14.2 71.0 713 5 ABAV45873 Aabv45873 Human pro
C 494 14.2 71.0 734 8 ACC78311 Acc78311 DNA encod
C 495 14.2 71.0 739 8 ACC78312 Acc78312 DNA encod
C 496 14.2 71.0 774 5 AAH64858 Aah64858 Human sec
C 497 14.2 71.0 779 4 AAL21741 Aaal21741 Human bre
C 498 14.2 71.0 783 11 ABD16810 Abdi16810 Pseudomon
C 499 14.2 71.0 801 11 ABD16872 Abdi16872 Pseudomon
C 500 14.2 71.0 807 4 AAH98307 Aah98307 Human EST

ALIGNMENTS

RESULT 1

AAAF33173
ID AAF33173 standard; DNA; 20 BP.
XX
AC AAF33173;
XX
DT 23-MAR-2001 (first entry)
XX
DE Human B7-1 antisense oligonucleotide SEQ ID NO: 256.
KW Human; mouse; B7-1; B7-2; antisense; PCR primer; inflammation;
KW autoimmune disorder; phosphorothioate backbone; ss.
XX
OS Homo sapiens.
XX
PN WO200074687-A1.
XX
PD 14-DEC-2000.
XX
PF 25-MAY-2000; 2000WO-US014471.
XX
PR 04-JUN-1999; 99US-00326186.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Vickers TA, Karras JG;

XX WPI; 2001-049991/06.
 XX Novel compound for diagnosing, preventing and treating immune disorders,
 PT comprising an oligonucleotide that specifically hybridizes with a nucleic
 PT acid sequence encoding B7 protein.
 XX Claim 11; Page 101; 162pp; English.
 PS
 XX The present invention provides sequences of antisense oligonucleotides
 CC targeted at the murine and human B7-1 and B7-2 coding and mRNA sequences.
 CC The antisense sequences have phosphorothioate backbones and some
 CC nucleotides are 2'-methoxyethoxy residues. The sequences can be used in
 CC the treatment of inflammatory and autoimmune disorders, including asthma,
 CC juvenile diabetes mellitus, myasthenia gravis, Graves' disease,
 CC rheumatoid arthritis, allograft rejection, inflammatory bowel disease,
 CC multiple sclerosis, psoriasis, systemic lupus erythematosus, contact
 CC dermatitis, rhinitis, allergies and cancer
 XX
 SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGTGTGTCGTGCTAGTCCC 20
 Db 1 CGTGTGTCGTGCTAGTCCC 20
 RESULT 2
 ADE27993
 ID ADE27993 standard; DNA; 20 BP.
 AC ADE27993;
 DT 29-JAN-2004 (first entry)
 XX Human B7-2 targeted oligonucleotide SEQ ID 255.
 DE ss; human; B7-2; inflammatory skin disorder; antisense; psoriasis;
 KW contact dermatitis; atopic dermatitis; seborrheic dermatitis;
 KW nummular dermatitis; generalised exfoliative dermatitis; eczema;
 KW critical costimulatory molecule.
 XX Synthetic.
 OS Homo sapiens.
 OS US2003176374-A1.
 PN 18-SEP-2003.
 PD 09-MAY-2001; 2001US-00851871.
 PF 31-DEC-1996; 96US-00777266.
 PR 04-JUN-1999; 99US-00326186.
 PR 25-MAY-2000; 2000WO-US014471.
 XX (BENN/) BENNETT C F.
 PA (VICK/) VICKERS T A.
 PA (KARR/) KARRAS J G.
 XX Bennett CF, Vickers TA, Karras JG;
 XX WPI; 2003-863863/80.
 DR Treating an inflammatory skin disorder such as psoriasis comprises
 PT topically applying an antisense compound targeted to the nucleic acid
 PT encoding human B7 protein.
 XX Example 20; SEQ ID NO 255; 88pp; English.
 PS The invention relates to a method of treating an inflammatory skin
 CC

CC disorder in an individual by topically applying an antisense compound
 CC targeted to a nucleic acid molecule encoding a human B7 protein. The
 CC invention is for treating an inflammatory skin disorder in individual.
 CC The skin disorder is psoriasis, contact dermatitis, atopic dermatitis,
 CC seborrheic dermatitis, nummular dermatitis, generalised exfoliative
 CC dermatitis or eczema. The invention effectively modulates critical
 CC costimulatory molecules such as the B7 protein. The present sequence
 CC represents a human B7-2 targeted oligonucleotide.
 XX
 SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 10; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGTGTGTCGTGCTAGTCCC 20
 Db 1 CGTGTGTCGTGCTAGTCCC 20
 RESULT 3
 ADJ54435
 ID ADJ54435 standard; DNA; 20 BP.
 XX ADJ54435;
 AC ADJ54435;
 DT 06-MAY-2004 (first entry)
 XX Human B7-2 DNA antisense oligonucleotide #63.
 DE Airway hyperresponsiveness; pulmonary inflammation;
 KW antisense oligonucleotide; human; B7 protein; B7-2; asthma;
 KW antiasthmatic; antiinflammatory; ss.
 XX Homo sapiens.
 OS US2004023917-A1.
 PN 05-FEB-2004.
 PD 23-MAY-2003; 2003US-00444206.
 PF 31-DEC-1996; 96US-00777266.
 PR 04-JUN-1999; 99US-00326186.
 PR 25-MAY-2000; 2000WO-US014471.
 PR 09-MAY-2001; 2001US-00851871.
 XX (BENN/) BENNETT C F.
 PA (VICK/) VICKERS T A.
 PA (KARR/) KARRAS J G.
 XX Bennett CF, Vickers TA, Karras JG;
 XX WPI; 2004-132608/13.
 DR Treating airway hyperresponsiveness or pulmonary inflammation comprises
 PT administering an antisense compound targeted to a nucleic acid molecule
 PT encoding a human B7 protein to the individual.
 XX Example 20; SEQ ID NO 255; 182pp; English.
 PS The invention relates to a method for treating airway hyperresponsiveness
 CC or pulmonary inflammation in an individual comprising administering an
 CC antisense compound targeted to a nucleic acid molecule encoding a human
 CC B7 protein. The invention also relates to a method of inhibiting
 CC expression of a nucleic acid molecule encoding B7-1 or B7-2. The
 CC antisense compound is an antisense oligonucleotide which has a modified
 CC sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2
 CC protein or both. The compound is useful for treating airway
 CC hyperresponsiveness or pulmonary inflammation, which is associated with
 CC asthma, by inhibiting expression of human B7 protein. This sequence
 CC represents an antisense oligonucleotide of the invention.
 XX

```
SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 1 CGTGTGCTGTGCTAGTCCC 20

RESULT 4
AA01045/c
ID AA01045 standard; cDNA; 124 BP.
XX
AC AA01045;
XX
DT 07-MAY-1996 (first entry)
XX
DE Human B7-2 exon h1A (signal).
XX
KW T-cell costimulatory molecule; B7-2; T-lymphocyte; CD28; CTLA4; receptor;
KW immunoglobulin; signal peptide; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 107..124
FT /*tag= a
XX
PN W09523859-A2.
XX
PD 08-SEP-1995.
XX
PF 02-MAR-1995; 95WO-US002576.
XX
PR 02-MAR-1994; 94US-00205697.
XX
PA (BGM ) BRIGHAM & WOMENS HOSPITAL.
PA (DAND ) DANA FARBER CANCER INST.
XX
PI Sharpe AH, Borriello F, Freeman GJ, Nadler LM;
DR WPI; 1995-320574/41.
DR P-PSDB; AAR82898.
XX
PT Novel T cell co-stimulatory molecules - corresponding to naturally
PT occurring alternatively spliced forms of T cells co-stimulatory molecules
or variants.
XX
PS Claim 40; Page 85; 111pp; English.
XX
CC Human T-cell costimulatory molecule B7-2 gene exon h1A (AA01045) encodes
CC the B7-2 signal peptide (AAR82898). Exon h1A can be utilised in the
CC construction of nucleic acids used in the prodn. of soluble forms of T-
CC cell costimulatory molecules that bind to CD28 or CTLA4 and trigger a
CC costimulatory signal in T-cells
XX
SQ Sequence 124 BP; 35 A; 30 C; 33 G; 26 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 5
AAC00427/c
ID AAC00427 standard; cDNA; 430 BP.
XX
```

```
AC AAC00427;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 425.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00421.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 425; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX
SQ Sequence 430 BP; 117 A; 94 C; 104 G; 113 T; 0 U; 2 Other;
Query Match 100.0%; Score 20; DB 3; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
   |||||
Db 75 CGTGTGCTGTGCTAGTCCC 56

RESULT 6
AAQ81351/c
ID AAQ81351 standard; cDNA; 1120 BP.
XX
AC AAQ81351;
XX
DT 25-MAR-2003 (revised)
DT 20-AUG-1995 (first entry)
XX
DE Human B lymphocyte antigen B7-2 (hB7-2-clone 29).
XX
KW CTLA4/CD28; counter receptor; B lymphocyte antigen; B7-2; ss.
XX
OS Homo sapiens.
XX
```

```

FH Key      Location/Qualifiers
FT CDS      107..1093
FT          /*tag= a
EN
PN
XX WO9503408-A1.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US008423.
XX
XX 26-JUL-1993; 93US-00101624.
XX
XX 19-AUG-1993; 93US-00109393.
XX
XX 03-NOV-1993; 93US-00147773.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX Freeman GJ, Nadler LM, Gray GS, Greenfield E;
XX
XX WPI; 1995-075236/10.
XX P-PSDB; AAR67984.
XX
XX Nucleic acids encoding CTLA4/CD28 counter receptor, B7-2 - useful for
XX enhancing or suppressing T-cell mediated immune responses.
XX
XX Claim 4; Fig 8; 175pp; English.
XX
XX A cDNA library was constructed in the pCDM8 vector using poly A+ RNA from
XX the human anti-IgM activated B cells. Four clones were strongly positive
XX for B7-2 expression by indirect immunofluorescence using CTLA4lg and flow
XX cytometric analysis. The B7-2 cDNA insert in clone 29 was sequenced in
XX the pCDM8 expression vector employing the following strategy. Initial
XX sequencing was performed using sequencing primers T7 (AAQ81352), CDM8R
XX (AAQ81353) (Invitrogen) homologous to pCDM8 vector sequences adjacent to
XX the clone B7-2 cDNA. Sequencing was performed using dye terminator
XX chemistry and an ABI automated DNA sequencer. DNA sequence obtd. using
XX these primers was used to design additional sequencing primers (see
XX AAQ81354-Q81363). This cycle of sequencing and selection of additional
XX primers was continued until the B7-2 cDNA was completely sequenced on
XX both strands. The human B7-2 clone 29 cDNA sequence is given in AAQ81351.
XX The predicted protein sequence (AAR67984) exhibits many features common
XX to other type 1 Ig superfamily membrane proteins. Following cleavage of
XX the signal peptide the resulting membrane-bound protein would have an
XX unmodified mol. wt. of approx. 34 kDa. The extracellular domain contains
XX eight potential N-linked glycosylation sites. E. coli transfected with a
XX vector contg. the cDNA insert of clone 29 was deposited under ATCC 69357
XX on July 26 1993. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGTCCC 20
Db |||||
|||
57 CGTGTGTCGTGCTAGTCCC 38

RESULT 7
AAAT49181/c
ID AAT49181 standard; cDNA; 1120 BP.
XX
XX AAT49181;
AC
XX
XX 08-APR-1997 (first entry)
DT
XX
XX Human B lymphocyte antigen B7-2 cDNA.
DE
XX
XX CTLA4; CD28; ligand; B7-2; B lymphocyte antigen; B-cell; costimulation;
XX immunoglobulin; antibody; autoimmune disease; allergy; tumour; vaccine;
XX graft versus host disease; T-cell; T lymphocyte; TH2 response;
XX immunosuppressive; immunostimulant; therapy; ss.

FH Key      Location/Qualifiers
FT CDS      107..1093
FT          /*tag= a
EN
PN
XX WO9503408-A1.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US008423.
XX
XX 26-JUL-1993; 93US-00101624.
XX
XX 19-AUG-1993; 93US-00109393.
XX
XX 03-NOV-1993; 93US-00147773.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX Freeman GJ, Nadler LM, Gray GS, Greenfield E;
XX
XX WPI; 1995-075236/10.
XX P-PSDB; AAR67984.
XX
XX Nucleic acids encoding CTLA4/CD28 counter receptor, B7-2 - useful for
XX enhancing or suppressing T-cell mediated immune responses.
XX
XX Claim 4; Fig 8; 175pp; English.
XX
XX A cDNA library was constructed in the pCDM8 vector using poly A+ RNA from
XX the human anti-IgM activated B cells. Four clones were strongly positive
XX for B7-2 expression by indirect immunofluorescence using CTLA4lg and flow
XX cytometric analysis. The B7-2 cDNA insert in clone 29 was sequenced in
XX the pCDM8 expression vector employing the following strategy. Initial
XX sequencing was performed using sequencing primers T7 (AAQ81352), CDM8R
XX (AAQ81353) (Invitrogen) homologous to pCDM8 vector sequences adjacent to
XX the clone B7-2 cDNA. Sequencing was performed using dye terminator
XX chemistry and an ABI automated DNA sequencer. DNA sequence obtd. using
XX these primers was used to design additional sequencing primers (see
XX AAQ81354-Q81363). This cycle of sequencing and selection of additional
XX primers was continued until the B7-2 cDNA was completely sequenced on
XX both strands. The human B7-2 clone 29 cDNA sequence is given in AAQ81351.
XX The predicted protein sequence (AAR67984) exhibits many features common
XX to other type 1 Ig superfamily membrane proteins. Following cleavage of
XX the signal peptide the resulting membrane-bound protein would have an
XX unmodified mol. wt. of approx. 34 kDa. The extracellular domain contains
XX eight potential N-linked glycosylation sites. E. coli transfected with a
XX vector contg. the cDNA insert of clone 29 was deposited under ATCC 69357
XX on July 26 1993. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGTCCC 20
Db |||||
|||
57 CGTGTGTCGTGCTAGTCCC 38

RESULT 8
AAV55784/c
ID AAV55784 standard; cDNA; 1120 BP.
XX
XX AAV55784;
AC
XX
XX 23-MAR-1999 (first entry)
DT
XX
XX Human B7-2 antigen coding sequence.
DE
XX
XX B7-2 antigen; mammalian tumour cell; T cell costimulation; CD28 ligand;
XX CTLA4 ligand; therapy; T-cell response; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key      Location/Qualifiers
FT CDS      107..1096
FT          /*tag= a

```

```

XX Homo sapiens.
XX
XX Key      Location/Qualifiers
FT CDS      107..1096
FT          /*tag= a
FT sig_peptide 107..175
FT          /*tag= b
FT mat_peptide 176..1093
FT          /*tag= c
XX
XX WO9640915-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-US009052.
XX
XX 07-JUN-1995; 95US-00479744.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX Freeman GJ, Nadler LM, Gray GS;
XX
XX WPI; 1997-077269/07.
XX P-PSDB; AAW08467.
XX
XX DNA encoding a B7-2 fusion protein - used to enhance or down regulate B
XX lymphocyte antigens.
XX
XX Example 4; Page 93-94; 171pp; English.
XX
XX A cDNA clone (AAT49181), designated clone 29, codes for the B-lymphocyte
XX antigen B7-2 (AAW08467), a CTLA4/CD28 ligand which costimulates T cell
XX activation. It was obtd. by transfecting COS cells with a human anti-IgM
XX activated B cell cDNA library, reacting transfectants with CTLA4lg and
XX CD28lg, and panning with anti-human IgG antibody. E. coli transfected
XX with a vector contg. the cDNA insert of clone 29 was deposited as ATCC
XX 69357. Nucleic acids encoding the extracellular domain, variable region-
XX like domain or constant region-like domain of B7-2 (see also AAT49197-98)
XX are used to construct novel fusion proteins with e.g. an immunoglobulin
XX constant region. These can be expressed in host cells and used to enhance
XX or suppress T cell-mediated immune responses
XX
XX Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 20; DB 2; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTCGTGCTAGTCCC 20
Db |||||
|||
57 CGTGTGTCGTGCTAGTCCC 38

RESULT 8
AAV55784/c
ID AAV55784 standard; cDNA; 1120 BP.
XX
XX AAV55784;
AC
XX
XX 23-MAR-1999 (first entry)
DT
XX
XX Human B7-2 antigen coding sequence.
DE
XX
XX B7-2 antigen; mammalian tumour cell; T cell costimulation; CD28 ligand;
XX CTLA4 ligand; therapy; T-cell response; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key      Location/Qualifiers
FT CDS      107..1096
FT          /*tag= a

```

```

XX US5861310-A.
XX
XX
XX PD 19-JAN-1999.
XX
XX PF 30-MAY-1995; 95US-00456104.
XX
XX PR 03-NOV-1993; 93US-00147773.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX PI Gray GS, Freeman GJ, Nadler LM;
XX
XX WPI; 1999-130394/11.
XX
XX DR P-PSDB; AAW73638.
XX
XX
XX PT Tumour cell transfected to express B7-2 molecule - useful for tumour
XX therapy by stimulating T-cell response.
XX
XX PS Claim 9; Col 27-30; 27pp; English.
XX
XX CC This sequence encodes the human B7-2 antigen, which can be used in the
XX method of the invention. The method is for transfecting an isolated
XX mammalian tumour cell with an exogenous nucleic acid molecule encoding a
XX mammalian B7-2 molecule, where the B7-2 molecule is expressed in the
XX tumour cell is capable of costimulating a T cell and is capable of
XX binding a CD28 or CTLA4 ligand. The method is useful for treating tumours
XX by stimulating a T-cell response against tumour cells in vivo
XX
XX SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 2; Length 1120;
XX Best Local Similarity 100.0%; Pred. No. 4;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CGTGTGCTGTGCTAGTCCC 20
XX |||||
XX Db 57 CGTGTGCTGTGCTAGTCCC 38
XX
XX RESULT 9
XX AAC84049/c
XX ID AAC84049 standard; cDNA; 1120 BP.
XX
XX AC AAC84049;
XX
XX DT 28-MAR-2001 (first entry)
XX
XX DE Human B lymphocyte antigen B7-2 cDNA clone 29.
XX
XX KW Immunomodulator; fusion protein; human; murine; mouse; lymphocyte; CD28;
XX antigen; extracellular domain; CTLA4; immunoglobulin constant region;
XX immunogenicity; tumour; sarcoma; antigen presenting cell; macrophage;
XX T cell-mediated immune response; transplantation; vaccination; ss.
XX
XX OS Homo sapiens.
XX
XX XN US6130316-A.
XX
XX PD 10-OCT-2000.
XX
XX PF 26-JUL-1994; 94US-00280757.
XX
XX PR 26-JUL-1993; 93US-00101624.
XX
XX PR 19-AUG-1993; 93US-00109393.
XX
XX PR 03-NOV-1993; 93US-00147773.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX (REPK ) REPLIGEN CORP.
XX
XX PI Freeman GJ, Nadler LM, Gray GS, Greenfield E;
XX
XX WPI; 2000-655681/63.

```

```

DR P-PSDB; AAB37085.
XX
XX Nucleic acids and fusion proteins of CTLA4/CD28 ligands, useful for
XX enhancing or suppressing T cell-mediated immune responses, especially
XX during tissue, skin or organ transplantation, or in graft-versus-host
XX disease.
XX
XX PS Claim 43; Fig 8; 83pp; English.
XX
XX CC The invention relates to an isolated nucleic acid molecule encoding a
XX fusion protein comprising a first nucleotide sequence encoding a first
XX peptide, and a second nucleotide sequence encoding a second peptide. The
XX first nucleotide sequence hybridizes in 6 x sodium chloride/sodium
XX citrate (SSC) at 45 deg. C, followed by a wash in 0.2 x SSC at 50 deg. C
XX to a portion of a nucleotide sequence which encodes a human or murine B
XX lymphocyte antigen (B7-2) extracellular domain. The first peptide has the
XX ability to bind CD28 or CTLA4. The first peptide has an amino acid
XX sequence that is identical or at least 50% identical with the
XX extracellular domain of a human B7-2 peptide (AAB37085). The second
XX peptide is especially an immunoglobulin constant region. This sequence
XX represents the cDNA clone 29 encoding the human B lymphocyte antigen B7-2
XX (hB7-2 clone 29) and is used as a first sequence in the construct of the
XX invention. The nucleic acid molecules are useful in various expression
XX vectors to direct synthesis of the corresponding proteins or peptides in
XX a variety of hosts, particularly eukaryotic cells, e.g. mammalian or
XX insect cell culture. The nucleic acids are also useful for enhancing the
XX immunogenicity of a mammalian cell, e.g. tumour cell (sarcoma) or an
XX antigen presenting cell (macrophage). The fusion proteins or peptides are
XX useful for enhancing or suppressing T cell-mediated immune responses,
XX e.g. in situations of tissue, skin or organ transplantation, or in graft-
XX versus-host disease. The proteins are also useful for enhancing the
XX efficacy of vaccination against a variety of pathogens, and may also be
XX used to upregulate an immune response against a particular pathogen
XX during an infection or against a tumour in a tumour-bearing host
XX
XX SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 3; Length 1120;
XX Best Local Similarity 100.0%; Pred. No. 4;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CGTGTGCTGTGCTAGTCCC 20
XX |||||
XX Db 57 CGTGTGCTGTGCTAGTCCC 38
XX
XX RESULT 10
XX ABV72340/c
XX ID ABV72340 standard; cDNA; 1120 BP.
XX
XX AC ABV72340;
XX
XX DT 16-DEC-2002 (first entry)
XX
XX DE Nucleotide sequence of human B7-2 (CD86).
XX
XX KW B7 protein; B7-2; CD86; CD28 ligand; T cell; T cell proliferation;
XX infectious disease; cancer; immunotherapy; immunotherapy; gene; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 107..1096
XX FT /*tag= a
XX FT /product= "B7-2"
XX
XX XN US2002115214-A1.
XX
XX PD 22-AUG-2002.
XX
XX PF 26-JAN-1996; 96US-00592711.
XX
XX PR 23-NOV-1988; 88US-00275433.

```

PR 07-APR-1992; 92US-00864805.
 PR 07-APR-1992; 92US-00864807.
 PR 07-APR-1992; 92US-00864866.
 PR 04-JUN-1993; 93US-00073223.
 PR 03-JUN-1994; 94US-00253964.
 PR 10-MAR-1995; 95US-00403253.
 PR 04-MAY-1995; 95US-00435816.
 XX (JUNE/) JUNE C H.
 PA (THOM/) THOMPSON C B.
 PA (NABEL/) NABEL G J.
 PA (GRAY/) GRAY G S.
 PA (RENN/) RENNETT P D.
 XX
 PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;
 XX
 DR WPI: 2002-712476/77.
 DR P-PSDB; ABB78364.
 XX
 PT Inducing a population of T cells to proliferate, by activating population
 PT of T cells and stimulating an accessory molecule on the surface of the T
 PT cells with a ligand which binds the accessory molecule.
 XX
 PS Disclosure; Page 41-43; 88pp; English.
 XX
 CC The present sequence encodes a member of the B7 family of protein, B7-2
 CC (CD86). B7 proteins are ligands for CD28. Activated T cells are contacted
 CC with a stimulatory form of a natural ligand for CD28, such as a B7
 CC protein, for costimulation. B7 molecules are used in the method of the
 CC invention. The specification describes method for inducing a population
 CC of T cells to proliferate. The method involves activating population of T
 CC cells, stimulating an accessory molecule (e.g. CD28) on T cell surface
 CC with a ligand (e.g. B7 protein) which binds the molecule, to induce
 CC proliferation of T cells, monitoring proliferation of T cells in response
 CC to continuing exposure to the ligand, and reactivating and restimulating
 CC T cells when rate of proliferation has decreased to induce further
 CC proliferation of the cells. The method is useful for inducing
 CC proliferation of T cells, for use in treatment of infectious disease,
 CC cancer and immunotherapy. The method allows for the expansion of a
 CC population of T cells in numbers sufficient to reconstitute an
 CC individual's total CD4+ or CD8+ T cell population. The resulting T cell
 CC population can be genetically transduced and used for immunotherapy or
 CC can be used in methods of in vitro analyses of infectious agents. A
 CC population of tumour-infiltrating lymphocytes can be obtained from an
 CC individual afflicted with cancer and the T cells stimulated to
 CC proliferate to sufficient numbers. The resulting T cell population can be
 CC genetically transduced to express tumour necrosis factor (TNF) or other
 CC factor and restored to the individual. CD4+ T cells expanded by this
 CC method are useful in the treatment of HIV infection in an individual
 XX
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 6; Length 1120;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGTGTGCTGTGCTAGTCCC 20
 DB 57 CGTGTGCTGTGCTAGTCCC 38
 RESULT 11
 AAD27968/c
 ID AAD27968 standard; cDNA; 1120 BP.
 XX
 AC AAD27968;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Human B7-2 cDNA.
 XX
 KW T cell; CD3; accessory molecule; CD28; cancer; infectious disease;
 immunotherapy; human immunodeficiency virus; HIV infection; cytokine;

KW human; B7-2; CD86; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 107..1096
 FT /*tag= a
 FT /product= "B7-2 protein"
 XX
 PN US6352694-B1.
 XX
 PD 05-MAR-2002.
 XX
 PF 10-MAR-1995; 95US-00403253.
 XX
 PR 03-JUN-1994; 94US-00253964.
 XX
 PA (GENY) GENETICS INST INC.
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;
 XX
 DR WPI: 2002-314696/35.
 DR P-PSDB; AAE14634.
 XX
 PT Inducing T cell population to proliferate, useful in cancer therapy,
 PT comprises activating T cells by contacting T cells in vitro with
 PT immobilized anti-CD3 antibody and stimulating accessory molecule on T
 PT cell surface.
 XX
 PS Example 11; Col 61-64; 71pp; English.
 CC The invention relates to a method of inducing T cell population to
 CC proliferate for use in therapy comprising activating T cells by
 CC contacting T cells in vitro with anti-CD3 antibody which is immobilised
 CC on solid phase surface, and stimulating accessory molecule on T cell
 CC surface in vitro with anti-CD28 antibody, or stimulatory form of natural
 CC ligand for CD28 such as B7-1 or B7-2. The method is useful for inducing a
 CC population of T cells to proliferate in sufficient numbers for use in
 CC therapy e.g., for treating cancer or an infectious disease. The method
 CC can be used to selectively expand the population of CD28+, CD4+, CD8+,
 CC CD28RA+ or CD28RO+ T cells for immunotherapy. The T cell population
 CC resulting by the method can be genetically transduced and used for
 CC immunotherapy or can be used for in vitro analysis of infectious agents
 CC such as human immunodeficiency virus (HIV). Proliferation of a population
 CC of CD4+ T cells obtained from an individual infected with HIV can be
 CC achieved and the cells rendered resistant to HIV infection. Following the
 CC expansion of the T cells to sufficient numbers, the expanded T cells are
 CC restored to the individual. Also CD4+ T cells expanded by the above
 CC mentioned is useful for treating HIV infection in an individual. A
 CC population of tumour-infiltrating lymphocytes can be obtained from an
 CC individual afflicted with cancer and the T cells stimulated to
 CC proliferate to sufficient numbers and restored to the individual. The
 CC supernatants from cultures of T cells expanded from above mentioned
 CC method are useful as a rich source of cytokines and can be used to
 CC sustain T cells in vivo or ex vivo. Stimulating and expanding a
 CC population of antigen specific T cells are useful in therapeutic
 CC conditions where it is desirable to upregulate an immune response. The T
 CC cell proliferation occurs in the absence of exogenous growth factors or
 CC accessory cells. The present sequence is human B7-2 (CD86) cDNA used in
 CC the invention
 XX
 SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 6; Length 1120;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGTGTGCTGTGCTAGTCCC 20
 DB 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 12
AAD60974/c
ID AAD60974 standard; cDNA; 1120 BP.
XX
AC AAD60974;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human CD28 ligand, B7-2 (CD86) cDNA.
XX
KW Human; HIV infection; human immunodeficiency virus; CD28 ligand; B7;
XX therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 107..1096
FT /*tag= a
FT /product= "CD28 ligand, B7-2 protein"
XX
PN US2003099643-A1.
XX
PD 29-MAY-2003.
XX
PF 08-JUL-1999; 99US-00350202.
XX
PR 23-NOV-1988; 88US-00275433.
PR 22-NOV-1989; 89WO-US005304.
PR 07-APR-1992; 92US-00864805.
PR 07-APR-1992; 92US-00864807.
PR 07-APR-1992; 92US-00864866.
PR 04-JUN-1993; 93US-00073223.
PR 03-JUN-1994; 94US-00253694.
PR 10-MAR-1995; 95US-00403253.
XX
(JUNE/) JUNE C H.
PA (THOM/) THOMPSON C B.
PA (NABE/) NABEL G J.
PA (GRAY/) GRAY G S.
PA (RENN/) RENNERT P D.
XX
FI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;
XX
DR WPI; 2003-801206/75.
DR P-PSDB; ABW00432.
XX
PT Treating HIV infection in individual by isolating T cells from
PT leukocytes, contacting T cells with anti-CD3 antibody for T cell
PT proliferation, separating antibody from T cells, monitoring proliferation
PT of T cells.
XX
PS Disclosure; Page 34-35; 0pp; English.
XX
CC The present invention relates to a novel method of treating human
CC immunodeficiency virus (HIV) infection in an individual. The method
CC involves isolating population of CD4 T cells from leukocytes, contacting
CC population of CD4+ T cells with an anti-CD3 antibody for stimulating T
CC cell proliferation, separating antibody from T cells, monitoring
CC proliferation of T cells, restimulating T cells with antibody and
CC restoring T cells to individual. The present sequence is human CD28
CC ligand, B7 cDNA. This sequence is used to illustrate the method of the
CC invention
XX
SQ Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
|||||
DB 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 13
ADK69859/c
ID ADK69859 standard; cDNA; 1120 BP.
XX
AC ADK69859;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human B7-2 cDNA.
XX
KW CD28-associated signal; immunotherapy; infectious disease; cancer;
XX leukopheresis; human; B7-2; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 107..1093
FT /*tag= a
FT /product= "Human B7-2 protein"
XX
PN US6534055-B1.
XX
PD 18-MAR-2003.
XX
PF 04-MAY-1995; 95US-00435816.
XX
PR 07-APR-1992; 92US-00864805.
PR 07-APR-1992; 92US-00864807.
PR 07-APR-1992; 92US-00864866.
PR 04-JUN-1993; 93US-00073223.
PR 03-JUN-1994; 94US-00253964.
PR 10-MAR-1995; 95US-00403253.
XX
(GEMY) GENETICS INST INC.

June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;
WPI; 2003-531074/50.
P-PSDB; ADK69860.

Expanding T cell populations, useful for preparing renewable sources of T cells (e.g. CD4+) for immunotherapy or diagnostics, by stimulating a CD28 -associated signal on the surface of the cells with an anti-CD28 antibody, B7-1 or B7-2.

Disclosure; SEQ ID NO 3; 82pp; English.

The invention relates to a method for expanding a population of T cells to about 100-100000-fold over the original T cell population, or to about 10 log 1.0 to 12 log 1.2. The method comprises stimulating a CD28-associated signal on the surface of the T cells with agent comprising an anti-CD28 antibody, B7-1 or B7-2. The method is useful for inducing a population of T cells (e.g. CD4+, CD28+, CD8+, CD28RA+ or CD28RO+ T cells) to proliferate. The method is particularly useful for preparing a renewable source of CD4+ T cells. The expanded T cell population can be genetically transduced, and used for immunotherapy to treat a variety of human diseases (e.g. infectious diseases or cancer), or used in diagnostic protocols. T cells were obtained from leukopheresis of a normal donor, and purified with FICOLL density gradient centrifugation, followed by magnetic immunobead sorting. The present sequence is human B7 -2 cDNA.

Sequence 1120 BP; 354 A; 237 C; 230 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20
|||||
DB 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 14
ADI19363/c
ID ADI19363 standard; cDNA; 1120 BP.
XX
AC ADI19363;
XX
DT 15-APR-2004 (first entry)
XX
DE Human B7-2 cDNA, CD86.
XX
KW T cell; ss; immunotherapy; therapy; HIV infection; cancer;
KW infectious disease; cytostatic; antimicrobial; gene; human; chromosome 3.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 107..1096
FT /*tag= a
FT /product= "Human B7-2 protein"
XX
PN US2004001829-A1.
XX
PD 01-JAN-2004.
XX
PF 17-MAR-2003; 2003US-00390330.
XX
PR 23-NOV-1988; 88US-00275433.
PR 22-NOV-1989; 89WO-US005304.
PR 07-APR-1992; 92US-00844805.
PR 07-APR-1992; 92US-00864807.
PR 07-APR-1992; 92US-00864866.
PR 04-JUN-1993; 93US-00073223.
PR 03-JUN-1994; 94US-00253164.
PR 10-MAR-1995; 95US-00403253.
PR 04-MAY-1995; 95US-00435816.
XX
PA (JUNE/) JUNE C H.
PA (THOM/) THOMPSON C B.
PA (NABE/) NABEL G J.
PA (GRAY/) GRAY G S.
PA (RENN/) RENNERT P D.
XX
PI June CH, Thompson CB, Nabel GJ, Gray GS, Rennert PD;
XX
DR WPI; 2004-061648/06.
DR P-PSDB; ADI19364.

Inducing a population of T cells to proliferate, for immunotherapy or treating HIV infection, cancer or infectious disease, comprises activating a population of T cells and stimulating an accessory molecule on the surface of the T cells.
XX
PS Disclosure; SEQ ID NO 3; 80pp; English.
XX
CC The invention relates to a method for inducing a population of T cells to proliferate. The method comprising activating a population of T cells, and stimulating an accessory molecule on the surface of the T cells with a ligand that binds the accessory molecule. The invention is useful for immunotherapy, for treating HIV infection, cancer or infectious disease, or in diagnostic applications. The present sequence is human B7-2 cDNA. B7 gene is located on chromosome 3.
XX
SQ Sequence 1120 BP; 354 A; 238 C; 229 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15
ADJ54476/c
ID ADJ54476 standard; DNA; 2781 BP.
XX
AC ADJ54476;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human B7-2 DNA #2.
XX
KW Airway hyperresponsiveness; pulmonary inflammation;
KW antisense oligonucleotide; human; B7 protein; B7-2; asthma;
KW antiasthmatic; antiinflammatory; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 117..1106
FT /*tag= a
FT /product= "Human B7-2 #2"
XX
PN US2004023917-A1.
XX
PD 05-FEB-2004.
XX
PF 23-MAY-2003; 2003US-00444206.
XX
PR 31-DEC-1996; 96US-00777266.
PR 04-JUN-1999; 99US-00326186.
PR 25-MAY-2000; 2000WO-US014471.
PR 09-MAY-2001; 2001US-00851871.
XX
PA (BENN/) BENNETT C F.
PA (VICK/) VICKERS T A.
PA (KARR/) KARRAS J G.
XX
PI Bennett CF, Vickers TA, Karras JG;
XX
DR WPI; 2004-132608/13.
DR P-PSDB; ADJ54626.
DR GENBANK; BC040261.
XX
PT Treating airway hyperresponsiveness or pulmonary inflammation comprises administering an antisense compound targeted to a nucleic acid molecule encoding a human B7 protein to the individual.
XX
PS Example 27; SEQ ID NO 296; 182pp; English.
XX
CC The invention relates to a method for treating airway hyperresponsiveness or pulmonary inflammation in an individual comprising administering an antisense compound targeted to a nucleic acid molecule encoding a human B7 protein. The invention also relates to a method of inhibiting expression of a nucleic acid molecule encoding B7-1 or B7-2. The antisense compound is an antisense oligonucleotide which has a modified sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2 protein or both. The compound is useful for treating airway hyperresponsiveness or pulmonary inflammation, which is associated with asthma, by inhibiting expression of human B7 protein. This sequence represents DNA encoding a human B7-2 polypeptide of the invention.
XX
SQ Sequence 2781 BP; 887 A; 575 C; 597 G; 722 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 2781;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
|||||
Db 67 CGTGTGCTGTGCTAGTCCC 48

RESULT 16

ADP10461/c
ID ADP10461 standard; DNA; 2781 BP.
XX
AC ADP10461;
XX
DT 12-AUG-2004 (first entry)
XX
DE Reference mRNA sequences for marker #138.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS Homo sapiens.
XX
FN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00331831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 80; SEQ ID NO 470; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.
XX
SQ Sequence 2781 BP; 886 A; 575 C; 598 G; 722 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 12; Length 2781;
Best Local Similarity 100.0%; Pred. No. 4.6; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;
Qy 1 CGTGTCTCTGCTAGTCC 20
Db 67 CGTGTCTCTGCTAGTCC 48
RESULT 17
ADJ54477/c
ID ADJ54477 standard; DNA; 68001 BP.
XX
AC ADJ54477;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human B7-2 genomic DNA.
XX

KW Airway hyperresponsiveness; pulmonary inflammation;
KW antisense oligonucleotide; human; B7 protein; B7-2; asthma;
KW antiasthmatic; antiinflammatory; gene; ds.
XX
OS Homo sapiens.
XX
FN US2004023917-A1.
XX
PD 05-FEB-2004.
XX
PF 23-MAY-2003; 2003US-0044206.
XX
PR 31-DEC-1996; 96US-00777266.
PR 04-JUN-1999; 99US-00326186.
PR 25-MAY-2000; 2000WO-US014471.
PR 09-MAY-2001; 2001US-00851871.
XX
PA (BENN/) BENNETT C F.
PA (VICK/) VICKERS T A.
PA (KARR/) KARRAS J G.
XX
FI Bennett CF, Vickers TA, Karras JG;
XX
DR WPI; 2004-132608/13.
DR GENBANK; NT_005543.
XX
PT Treating airway hyperresponsiveness or pulmonary inflammation comprises
PT administering an antisense compound targeted to a nucleic acid molecule
PT encoding a human B7 protein to the individual.
XX
PS Example 27; SEQ ID NO 297; 182pp; English.
XX
CC The invention relates to a method for treating airway hyperresponsiveness
CC or pulmonary inflammation in an individual comprising administering an
CC antisense compound targeted to a nucleic acid molecule encoding a human
CC B7 protein. The invention also relates to a method of inhibiting
CC expression of a nucleic acid molecule encoding B7-1 or B7-2. The
CC antisense compound is an antisense oligonucleotide which has a modified
CC sugar moiety and nucleobase. The human B7 protein is human B7-1 or B7-2
CC protein or both. The compound is useful for treating airway
CC hyperresponsiveness or pulmonary inflammation, which is associated with
CC asthma, by inhibiting expression of human B7 protein. This sequence
CC represents human B7-2 genomic DNA of the invention.
XX
SQ Sequence 68001 BP; 20352 A; 13494 C; 14117 G; 20038 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 12; Length 68001;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;
Qy 1 CGTGTCTCTGCTAGTCC 20
Db 1067 CGTGTCTCTGCTAGTCC 1048
RESULT 18
ADN72456
ID ADN72456 standard; cDNA; 3168 BP.
XX
AC ADN72456;
XX
DT 15-JUL-2004 (first entry)
XX
DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 351.
XX
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS Arabidopsis thaliana.
XX
FN WO2004035798-A2.
XX

PD 29-APR-2004.
 XX
 PF 20-OCT-2003; 2003WO-EP011658.
 XX
 PR 18-OCT-2002; 2002EP-00079408.
 XX
 PA (CROP-) CROPDESIGN NV.
 XX
 PI Inze D, De Veylder L, Vlieghe K;
 XX
 DR WPI; 2004-348466/32.
 DR P-PSDB; ADM72457.
 XX
 XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprising modifying in a plant, expression of
 PT one or more nucleic acids and/or modifying level or activity of one or
 PT more proteins.
 XX
 PS Claim 1; SEQ ID NO 351; 134pp; English.
 XX
 CC This invention relates to a novel method for altering one or more plant
 CC characteristics. Specifically, it refers to identifying genes that are up
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 CC alter plant characteristics accordingly. The present invention describes
 CC generating transgenic plants for the production of growth regulators,
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 CC or physiology, altered endoreplication, biochemistry, signal
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,
 CC each relative to the corresponding wild type plants. Accordingly, these
 CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polynucleotide sequence is the cress cDNA
 CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
 CC transcription factor, given in an exemplification of the invention.
 XX
 SQ Sequence 3168 BP; 1042 A; 542 C; 760 G; 724 T; 0 U; 0 Other;
 Query Match 87.0%; Score 17.4; DB 12; Length 3168;
 Best Local Similarity 94.7%; Pred. No. 95;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GTGTCTCTGCTAGTCC 20
 Db 2429 GTGTCTCTGCTAGTCC 2447
 RESULT 19
 AAC05026
 ID AAC05026 standard; cDNA; 368 BP.
 XX
 AC AAC05026;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 9101.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 FN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 PR

XX (GEST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Giordano J;
 PI
 XX WPI; 2000-500381/45.
 DR
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 9101; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dr primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors
 XX
 SQ Sequence 368 BP; 79 A; 122 C; 59 G; 105 T; 0 U; 3 Other;
 Query Match 80.0%; Score 16; DB 3; Length 368;
 Best Local Similarity 88.9%; Pred. No. 3.6e+02;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GTGTCTCTGCTAGTCC 19
 Db 192 GTGTCTCTGCTAGTCC 209
 RESULT 20
 AAL01062
 ID AAL01062 standard; cDNA; 372 BP.
 XX
 AC AAL01062;
 XX
 DT 21-NOV-2001 (first entry)
 XX
 DE Human reproductive system related antigen cDNA SEQ ID NO: 1063.
 XX
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200155320-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001339.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225211P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
FI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
DR P-PSDB; AAM95092.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; SEQ ID NO 1063; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention
XX
SQ Sequence 372 BP; 104 A; 68 C; 96 G; 101 T; 0 U; 3 Other;
Query Match 80.0%; Score 16; DB 4; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;

	Matches	16;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	2	GTGTCCTCTGCTAGT	17							
Db	70	GTGTCCTCTGCTAGT	85							
RESULT 21										
ABL96529	ID ABL96529 standard; cDNA; 372 BP.									
XX	AC	ABL96529;								
XX	DT	21-JUN-2002 (first entry)								
DE	DE	Human testicular antigen encoding cDNA SEQ ID NO: 197.								
KW	KW	Human; testicular antigen; testes; cancer; metastasis; immune disorder;								
KW	KW	reproductive system disorder; urinary system disorder; gene therapy;								
KW	KW	cardiovascular disorder; respiratory disorder; neurological disorder;								
KW	KW	gastrointestinal disease; infection; cytostatic; gene; ss.								
OS	OS	Homo sapiens.								
XX	PN	WO200155317-A2.								
PD	PD	02-AUG-2001.								
XX	PF	17-JAN-2001; 2001WO-US001329.								
XX	PR	31-JAN-2000; 2000US-0179065P.								
PR	PR	04-FEB-2000; 2000US-0180628P.								
PR	PR	24-FEB-2000; 2000US-0184664P.								
PR	PR	02-MAR-2000; 2000US-0186350P.								
PR	PR	16-MAR-2000; 2000US-0189874P.								
PR	PR	17-MAR-2000; 2000US-0190076P.								
PR	PR	18-APR-2000; 2000US-0198123P.								
PR	PR	19-MAY-2000; 2000US-0205515P.								
PR	PR	07-JUN-2000; 2000US-0209467P.								
PR	PR	28-JUN-2000; 2000US-0214886P.								
PR	PR	30-JUN-2000; 2000US-0215135P.								
PR	PR	07-JUL-2000; 2000US-0216647P.								
PR	PR	07-JUL-2000; 2000US-0216880P.								
PR	PR	11-JUL-2000; 2000US-0217487P.								
PR	PR	11-JUL-2000; 2000US-0217496P.								
PR	PR	14-JUL-2000; 2000US-0218290P.								
PR	PR	26-JUL-2000; 2000US-0220963P.								
PR	PR	26-JUL-2000; 2000US-0220964P.								
PR	PR	14-AUG-2000; 2000US-0224518P.								
PR	PR	14-AUG-2000; 2000US-0225213P.								
PR	PR	14-AUG-2000; 2000US-0225214P.								
PR	PR	14-AUG-2000; 2000US-0225266P.								
PR	PR	14-AUG-2000; 2000US-0225267P.								
PR	PR	14-AUG-2000; 2000US-0225268P.								
PR	PR	14-AUG-2000; 2000US-0225270P.								
PR	PR	14-AUG-2000; 2000US-0225447P.								
PR	PR	14-AUG-2000; 2000US-0225757P.								
PR	PR	14-AUG-2000; 2000US-0225758P.								
PR	PR	14-AUG-2000; 2000US-0225759P.								
PR	PR	18-AUG-2000; 2000US-0226279P.								
PR	PR	22-AUG-2000; 2000US-0226681P.								
PR	PR	22-AUG-2000; 2000US-0226868P.								
PR	PR	22-AUG-2000; 2000US-0227182P.								
PR	PR	23-AUG-2000; 2000US-0227009P.								
PR	PR	30-AUG-2000; 2000US-0228924P.								
PR	PR	01-SEP-2000; 2000US-0229287P.								
PR	PR	01-SEP-2000; 2000US-0229343P.								
PR	PR	01-SEP-2000; 2000US-0229344P.								
PR	PR	01-SEP-2000; 2000US-0229345P.								
PR	PR	05-SEP-2000; 2000US-0229509P.								
PR	PR	05-SEP-2000; 2000US-0229513P.								
PR	PR	06-SEP-2000; 2000US-0230437P.								
PR	PR	06-SEP-2000; 2000US-0230438P.								
PR	PR	08-SEP-2000; 2000US-0231242P.								
PR	PR	08-SEP-2000; 2000US-0231243P.								
PR	PR	08-SEP-2000; 2000US-0231244P.								
PR	PR	08-SEP-2000; 2000US-0231413P.								
PR	PR	08-SEP-2000; 2000US-0231414P.								
PR	PR	08-SEP-2000; 2000US-0232080P.								
PR	PR	08-SEP-2000; 2000US-0232081P.								
PR	PR	12-SEP-2000; 2000US-0231968P.								
PR	PR	14-SEP-2000; 2000US-0232397P.								
PR	PR	14-SEP-2000; 2000US-0232398P.								
PR	PR	14-SEP-2000; 2000US-0232399P.								
PR	PR	14-SEP-2000; 2000US-0232400P.								
PR	PR	14-SEP-2000; 2000US-0232401P.								
PR	PR	14-SEP-2000; 2000US-0233063P.								
PR	PR	14-SEP-2000; 2000US-0233064P.								
PR	PR	14-SEP-2000; 2000US-0233065P.								
PR	PR	21-SEP-2000; 2000US-0234223P.								
PR	PR	21-SEP-2000; 2000US-0234274P.								
PR	PR	25-SEP-2000; 2000US-0234997P.								
PR	PR	25-SEP-2000; 2000US-0234998P.								
PR	PR	26-SEP-2000; 2000US-0235484P.								
PR	PR	27-SEP-2000; 2000US-0235834P.								
PR	PR	27-SEP-2000; 2000US-0235836P.								
PR	PR	29-SEP-2000; 2000US-0236327P.								
PR	PR	29-SEP-2000; 2000US-0236367P.								
PR	PR	29-SEP-2000; 2000US-0236368P.								
PR	PR	29-SEP-2000; 2000US-0236369P.								
PR	PR	29-SEP-2000; 2000US-0236370P.								
PR	PR	02-OCT-2000; 2000US-0236802P.								
PR	PR	02-OCT-2000; 2000US-0237037P.								
PR	PR	02-OCT-2000; 2000US-0237038P.								
PR	PR	02-OCT-2000; 2000US-0237039P.								
PR	PR	02-OCT-2000; 2000US-0237040P.								
PR	PR	13-OCT-2000; 2000US-0239935P.								
PR	PR	13-OCT-2000; 2000US-0239937P.								
PR	PR	20-OCT-2000; 2000US-0240960P.								
PR	PR	20-OCT-2000; 2000US-0241221P.								
PR	PR	20-OCT-2000; 2000US-0241785P.								
PR	PR	20-OCT-2000; 2000US-0241786P.								
PR	PR	20-OCT-2000; 2000US-0241787P.								
PR	PR	20-OCT-2000; 2000US-0241808P.								
PR	PR	20-OCT-2000; 2000US-0241809P.								
PR	PR	01-NOV-2000; 2000US-0241826P.								
PR	PR	01-NOV-2000; 2000US-0244617P.								
PR	PR	08-NOV-2000; 2000US-0246474P.								
PR	PR	08-NOV-2000; 2000US-0246475P.								
PR	PR	08-NOV-2000; 2000US-0246476P.								
PR	PR	08-NOV-2000; 2000US-0246477P.								
PR	PR	08-NOV-2000; 2000US-0246478P.								
PR	PR	08-NOV-2000; 2000US-0246523P.								
PR	PR	08-NOV-2000; 2000US-0246524P.								
PR	PR	08-NOV-2000; 2000US-0246525P.								
PR	PR	08-NOV-2000; 2000US-0246526P.								
PR	PR	08-NOV-2000; 2000US-0246527P.								
PR	PR	08-NOV-2000; 2000US-0246528P.								
PR	PR	08-NOV-2000; 2000US-0246532P.								
PR	PR	08-NOV-2000; 2000US-0246609P.								
PR	PR	08-NOV-2000; 2000US-0246610P.								
PR	PR	08-NOV-2000; 2000US-0246611P.								
PR	PR	08-NOV-2000; 2000US-0246613P.								
PR	PR	17-NOV-2000; 2000US-0249207P.								
PR	PR	17-NOV-2000; 2000US-0249208P.								
PR	PR	17-NOV-2000; 2000US-0249209P.								
PR	PR	17-NOV-2000; 2000US-0249210P.								
PR	PR	17-NOV-2000; 2000US-0249211P.								
PR	PR	17-NOV-2000; 2000US-0249212P.								
PR	PR	17-NOV-2000; 2000US-0249213P.								
PR	PR	17-NOV-2000; 2000US-0249214P.								
PR	PR	17-NOV-2000; 2000US-0249215P.								
PR	PR	17-NOV-2000; 2000US-0249216P.								
PR	PR	17-NOV-2000; 2000US-0249217P.								
PR	PR	17-NOV-2000; 2000US-0249218P.								

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251473P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259670P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483232/52.
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX Claim 1; SEQ ID NO 197; 766pp; English.
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a cDNA of the
CC invention
XX
XX Sequence 372 BP; 104 A; 68 C; 96 G; 101 T; 0 U; 3 Other;
SQ Query Match 80.0%; Score 16; DB 4; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGT 17
DB 70 GTGTGCTGTGCTAGT 85

RESULT 22
ABX91230
ID ABX91230 standard; cDNA; 515 BP.
XX
XX ABX91230;
XX
XX 07-MAY-2003 (first entry)
XX Murine gene trapped sequence (GTS) SEQ ID NO 574.
XX Murine; mouse; gene trap technology; gene trapped sequence; GTS;
KW gene identification; functional genomic analysis; gene discovery;
KW gene expression analysis; cross species hybridisation analysis;
KW antisense inhibition; gene targeting; gene; ss.
XX
XX Mus sp.
XX
XX US2002161207-A1.
XX
XX 31-OCT-2002.
XX
XX 30-NOV-2000; 2000US-00728444.
XX

XX 01-DEC-1999; 99US-0168360P.
XX (FRIE/) FRIEDRICH G.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.
XX Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2003-288124/28.
XX New murine polynucleotides comprising gene trapped sequences, useful in
PT functional genomic analysis, in the development of new therapeutic or
PT diagnostic agents, for diagnostic gene expression analysis or for genetic
PT manipulations.
XX
XX Claim 2; SEQ ID NO 574; 29pp; English.
XX The present invention relates to novel murine cDNAs produced using gene
CC trap technology. The OMNIBANK gene trapped sequences (GTSS) are
CC individually identified novel genes, and are useful in functional genomic
CC analysis, in the discovery and development of new therapeutic and
CC diagnostic agents, for gene discovery, for diagnostic gene expression
CC analysis, for cross species hybridisation analysis, and for genetic
CC manipulations such as antisense inhibition or gene targeting. The
CC polynucleotides of the invention are also useful for isolating cDNAs,
CC genomic clones or full-length genes/polynucleotides, or their homologues,
CC heterologues, paralogues or orthologues, that are capable of hybridising
CC to one or more of the new murine polynucleotide sequences. The
CC polynucleotides are also useful for identifying the coding regions of the
CC murine genome, and as hybridisation probes. ABX90657-ABX91862 represent
CC the murine GTSS of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/peipsDIBentry.html
XX
XX Sequence 515 BP; 152 A; 101 C; 109 G; 151 T; 0 U; 2 Other;
SQ Query Match 80.0%; Score 16; DB 8; Length 515;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAG 16
DB 384 CGTGTGCTGTGCTAG 399

RESULT 23
ABZ21938
ID ABZ21938 standard; cDNA; 2175 BP.
XX
XX ABZ21938;
XX
XX 28-MAR-2003 (first entry)
XX Human CAK1 antigen 9.46 encoding cDNA SEQ ID NO:1.
XX Human; CAK1 antigen 9.46; antigen; malignant tumour; inflammation;
KW immunological disease; haemopathy; HIV infection; gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
PH 245..505
FT CDS /*tag= a
FT /product= "CAK1 antigen 9.46"
XX
XX CN1352019-A.
XX
XX 05-JUN-2002.
XX
XX 10-NOV-2000; 2000CN-00127321.
XX

CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g. osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention

XX
 SQ Sequence 439 BP; 68 A; 116 C; 180 G; 74 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 439;
 Best Local Similarity 89.5%; Pred. No. 4.7e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCC 19
 |||||
 Db 39 CGTGTCTCTGCGAGGCC 57
 |||||

RESULT 26

AAK82836/c
 ID AAK82836 standard; DNA; 468 BP.

XX AAK82836;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37648.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 21-SEP-2000; 2000US-0234275P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.

PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244611P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246603P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-483426/52.	
XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PT	Disclosure; SEQ ID NO 37648; 3071pp + Sequence Listing; English.	
XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
XX	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	

CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX
SQ Sequence 468 BP; 93 A; 136 C; 148 G; 91 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 468;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps

Qy 2 GTGTGTCCTGCTAGTCC 20
| | | | | | | | | | | | | | | | | | | | | |
Db 53 GGCTGTCCTGGCTAGTCC 35

RESULT 27
AAD51678/c

ID AAD51678 standard; cDNA; 1345 BP.

XX
XX
AC AAD51678;
XX
XX
DT 16-APR-2003 (first entry)
XX
XX
DE Human nucleic acid associated protein (NAAP)-4 encoding cDNA.
XX
XX
KW Human; nucleic acid associated protein; NAAP; cancer; atherosclerosis;
KW cell proliferative disorder; neurological disorder; Huntington's disease;
KW epilepsy; stroke; immune disorder; acquired immune deficiency syndrome;
KW AIDS; inflammatory disorder; allergy; developmental disorder; infection;
KW hypothyroidism; cushing's syndrome; gene therapy; cytostatic; neutropic;
KW anticonvulsant; neuroprotective; cerebroprotective; thyromimetic; gene;
KW ss.
XX
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 301..1074
FT FT /*tag= a
FT FT /product= "Human NAAP-4 protein"
FT FT sig_peptide 301..420
FT FT /*tag= b
FT FT mat_peptide 421..1071
FT FT /*tag= c
FT FT /product= "Human mature NAAP-4 protein"
XX
XX
PN WO200299115-A2.
XX
XX
PD 12-DEC-2002.
XX
XX
PF 31-MAY-2002; 2002WO-US017050.
XX
XX
PR 01-JUN-2001; 2001US-0295359P.
PR 08-JUN-2001; 2001US-0296878P.
PR 08-JUN-2001; 2001US-0297222P.
PR 15-JUN-2001; 2001US-0298615P.
PR 15-JUN-2001; 2001US-0298665P.
PR 15-JUN-2001; 2001US-0298693P.
PR 21-JUN-2001; 2001US-0300176P.
PR 19-APR-2002; 2002US-0373891P.
XX
XX
PA (INCY-) INCYTE GENOMICS INC.
PA (YUEH/) YUE H.
XX
XX
PI Yue H, Tang YT, Baughn MR, Becha SD, Warren BA, Wallia NK;
PI Lal PG, Lee EA, Hafalla AJA, Richardson TW, Griffin JA, Emerling BM
PI Raskumar J, Yue H, Swarnakar A, Tran B, Li JX, Yao MG, Yang J;
PI Ison CH, Forsythe IJ, Honchell CD, Arvizu CS, Elliott VS, Lu Y;
PI Dong L, Luo W, Wang YE, Burford N, Borowsky ML, Nguyen DB;
PI Chinn AM, Kable AE;
XX

DR WPI; 2003-140626/13.
DR P-PSDB; AAE33764.
XX
PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
PT infections.
XX
PS Claim 89; Col 231; 257pp; English.
XX
CC The present invention relates to human nucleic acid associated proteins
CC (NAAP) and polynucleotides encoding such proteins. NAAP sequences are
CC useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of NAAP such
CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
CC disorders or infections. They are also used in gene therapy. The present
CC sequence is human NAAP-4 cDNA
XX
SQ Sequence 1345 BP; 255 A; 425 C; 456 G; 209 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 10; Length 1345;
Best Local Similarity 89.5%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGTGTGCTGCTAGTCC 19
DB 1134 CGTGTGCTGCTAGTCC 1116
RESULT 28
ABZ11458
ID ABZ11458 standard; cDNA; 2304 BP.
XX
AC ABZ11458;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 340.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Auandi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR P-PSDB; ABP69241.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.

XX
PS Claim 1; SEQ ID NO 340; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burn, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2304 BP; 384 A; 800 C; 720 G; 400 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 6; Length 2304;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGTGTGCTGCTAGTCC 19
DB 651 CGTGTGCTGCTAGTCC 669
RESULT 29
ADM43976
ID ADM43976 standard; cDNA; 2304 BP.
XX
AC ADM43976;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #340.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.
PR 20-AUG-2002; 2002US-00225251.
XX
PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
XX
PI Tang YT, Xue A, Drmanac RT;
XX
DR WPI; 2004-238579/22.
XX
PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
PS Disclosure; SEQ ID NO 340; 51pp; English.
XX
CC The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or

CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.

XX SQ Sequence 2304 BP; 384 A; 800 C; 720 G; 400 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 12; Length 2304;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCC 19
DB 651 CGTGTCTCTGCGAGGCC 669

RESULT 30
AAI66465
ID AAI66465 standard; cDNA; 2434 BP.
XX AC AAI66465;
XX DT 04-DEC-2001 (first entry)
XX DE P53 tumour antigen protein 13 coding sequence.
XX KW P53 tumour antigen protein 13; cancer; haemopathy; HIV infection;
XX immunological disease; inflammation; gene therapy; ss.
XX OS Unidentified.
XX PN CN1301754-A.
XX PD 04-JUL-2001.
XX PF 29-DEC-1999; 99CN-00127233.
XX PR 29-DEC-1999; 99CN-00127233.
XX PA (UYFU-) UNIV FUDAN.
XX PI Mao Y, Xie Y;
XX DR WPI; 2001-558148/63.
XX P-PSDB; AAG78691.
XX PT New polypeptide-P53 tumor antigen protein 13 and a polynucleotide
XX encoding the polypeptide.
XX PS Claim 6; Page 23-24(Disclosure); 31pp; Chinese.
XX CC The present invention provides the protein and coding sequences of P53
XX tumour antigen protein 13. The sequences can be used in the treatment of
XX cancer, haemopathy, HIV infection, immunological diseases and
XX inflammation. The present sequence is the coding sequence of the
XX invention

XX SQ Sequence 2434 BP; 562 A; 728 C; 568 G; 576 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 4; Length 2434;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20
DB 783 GTGGGTCTGTGCTTGTCCC 801

RESULT 31
ABL09670
ID ABL09670 standard; cDNA; 3545 BP.

XX ABL09670;
XX 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 23492.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKS) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX P-PSDB; ABB65567.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX PS Claim 1; SEQ ID NO 23492; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3545 BP; 966 A; 748 C; 742 G; 1089 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 4; Length 3545;
Best Local Similarity 89.5%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20
DB 2615 GTGTCTCTGCTAGTCCC 2633

RESULT 32
ADE55103
ID ADE55103 standard; DNA; 5315 BP.
XX AC ADE55103;
XX DT 29-JAN-2004 (first entry)
XX DE Rat gene AF059030, SEQ ID NO 908.
XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.

XX 27-FEB-2003.
PD 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; AF059030.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 5315 BP; 1349 A; 1324 C; 1288 G; 1354 T; 0 U; 0 Other;
SQ Query Match 79.0%; Score 15.8; DB 10; Length 5315;
Best Local Similarity 89.5%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GTGTGCTGTGCTAGTCCC 20
DB 3039 GTGGGTCGTGCTAGTCTC 3057
RESULT 33
ABL09674
ID ABL09674 standard; cDNA; 5856 BP.
XX ABL09674;
AC ABL09674;
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 23504.
DE Drosophila; developmental biology; cell signalling; insecticide;
XX Drosophila; developmental biology; cell signalling; insecticide;
XX

KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
FI WPI; 2001-656860/75.
DR P-PSDB; ABB65571.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
FT interactions.
XX Claim 1; SEQ ID NO 23504; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 5856 BP; 1527 A; 1429 C; 1445 G; 1455 T; 0 U; 0 Other;
SQ Query Match 79.0%; Score 15.8; DB 4; Length 5856;
Best Local Similarity 89.5%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GTGTGCTGTGCTAGTCCC 20
DB 421 GTGTGCTGTGCTAGTCCC 439
RESULT 34
ABX09140
ID ABX09140 standard; DNA; 82993 BP.
XX AC ABX09140;
XX 08-APR-2003 (first entry)
DT Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv221.
XX Mycobacterium tuberculosis.
XX Mycobacteriosis; survival; virulence; protective antigen; vaccine;
KW mycobacterial disease; tuberculosis; leprosy; ds; cosmid.
XX Mycobacterium tuberculosis.
XX WO200274903-A2.
XX 26-SEP-2002.
XX 22-FEB-2002; 2002WO-IB001973.
XX 22-FEB-2001; 2001US-0270123P.
XX (INSP) INST PASTEUR.
XX

PI Cole S;
XX
DR WPI; 2002-759885/82.
XX
PT Identifying and selecting genes for survival or virulence of mycobacteria
PT by a comparative genomic analysis of the sequences of Mycobacterium
PT tuberculosis and M. leprae.
XX
XX Disclosure; Fig 4; 874pp; English.
XX
XX This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a Mycobacterial cosmid DNA
CC sequence used in the method of the invention
XX

SQ Sequence 82993 BP; 14881 A; 2628 C; 26719 G; 14765 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 6; Length 82993;
Best Local Similarity 89.5%; Pred. No. 9.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20
DB 29325 GTGTGCTGTGCTGTCGC 29343
|||||

RESULT 35
AAI99682_21
Continuation (22 of 45) of AAI99682 from base 2100001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682
WP Fragment Name Begin End
WP AAI99682_00 1 110000
WP AAI99682_01 100001 210000
WP AAI99682_02 200001 310000
WP AAI99682_03 300001 410000
WP AAI99682_04 400001 510000
WP AAI99682_05 500001 610000
WP AAI99682_06 600001 710000
WP AAI99682_07 700001 810000
WP AAI99682_08 800001 910000
WP AAI99682_09 900001 1010000
WP AAI99682_10 1000001 1110000
WP AAI99682_11 1100001 1210000
WP AAI99682_12 1200001 1310000
WP AAI99682_13 1300001 1410000
WP AAI99682_14 1400001 1510000
WP AAI99682_15 1500001 1610000
WP AAI99682_16 1600001 1710000
WP AAI99682_17 1700001 1810000
WP AAI99682_18 1800001 1910000
WP AAI99682_19 1900001 2010000
WP AAI99682_20 2000001 2110000
WP AAI99682_21 2100001 2210000
WP AAI99682_22 2200001 2310000
WP AAI99682_23 2300001 2410000
WP AAI99682_24 2400001 2510000
WP AAI99682_25 2500001 2610000
WP AAI99682_26 2600001 2710000
WP AAI99682_27 2700001 2810000
WP AAI99682_28 2800001 2910000
WP AAI99682_29 2900001 3010000
WP AAI99682_30 3000001 3110000
WP AAI99682_31 3100001 3210000
WP AAI99682_32 3200001 3310000

WP AAI99682_33 3300001 3410000
WP AAI99682_34 3400001 3510000
WP AAI99682_35 3500001 3610000
WP AAI99682_36 3600001 3710000
WP AAI99682_37 3700001 3810000
WP AAI99682_38 3800001 3910000
WP AAI99682_39 3900001 4010000
WP AAI99682_40 4000001 4110000
WP AAI99682_41 4100001 4210000
WP AAI99682_42 4200001 4310000
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529

Query Match 79.0%; Score 15.8; DB 4; Length 110000;
Best Local Similarity 89.5%; Pred. No. 9.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20
DB 44936 GTGTGCTGTGCTGTCGC 44954
|||||

RESULT 36
AAI99683_21
Continuation (22 of 44) of AAI99683 from base 2100001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession Aai99683
WP Fragment Name Begin End
WP AAI99683_00 1 110000
WP AAI99683_01 100001 210000
WP AAI99683_02 200001 310000
WP AAI99683_03 300001 410000
WP AAI99683_04 400001 510000
WP AAI99683_05 500001 610000
WP AAI99683_06 600001 710000
WP AAI99683_07 700001 810000
WP AAI99683_08 800001 910000
WP AAI99683_09 900001 1010000
WP AAI99683_10 1000001 1110000
WP AAI99683_11 1100001 1210000
WP AAI99683_12 1200001 1310000
WP AAI99683_13 1300001 1410000
WP AAI99683_14 1400001 1510000
WP AAI99683_15 1500001 1610000
WP AAI99683_16 1600001 1710000
WP AAI99683_17 1700001 1810000
WP AAI99683_18 1800001 1910000
WP AAI99683_19 1900001 2010000
WP AAI99683_20 2000001 2110000
WP AAI99683_21 2100001 2210000
WP AAI99683_22 2200001 2310000
WP AAI99683_23 2300001 2410000
WP AAI99683_24 2400001 2510000
WP AAI99683_25 2500001 2610000
WP AAI99683_26 2600001 2710000
WP AAI99683_27 2700001 2810000
WP AAI99683_28 2800001 2910000
WP AAI99683_29 2900001 3010000
WP AAI99683_30 3000001 3110000
WP AAI99683_31 3100001 3210000
WP AAI99683_32 3200001 3310000
WP AAI99683_33 3300001 3410000
WP AAI99683_34 3400001 3510000
WP AAI99683_35 3500001 3610000
WP AAI99683_36 3600001 3710000
WP AAI99683_37 3700001 3810000
WP AAI99683_38 3800001 3910000
WP AAI99683_39 3900001 4010000
WP AAI99683_40 4000001 4110000
WP AAI99683_41 4100001 4210000
WP AAI99683_42 4200001 4310000
WP AAI99683_43 4300001 4403765

Query Match 79.0%; Score 15.8; DB 4; Length 110000;

Best Local Similarity 89.5%; Pred. NO. 9.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20
DB 42152 GTGTGCTGTGCTGTGCTGC 42170

RESULT 37
ADL13471
ID ADL13471 standard; DNA; 214520 BP.
XX AC ADL13471;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #3.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
OS WO2003054166-A2.
XX PN 03-JUL-2003.
XX PD 19-DEC-2002; 2002WO-US041225.
XX PF 20-DEC-2001; 2001US-0342603P.
XX PR (INCY-) INCYTE GENOMICS INC.
XX PA Jones KA, Schafer A;
XX PI WPI; 2003-559141/52.
XX DR Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PT a protein.
XX PS Disclosure; SEQ ID NO 3; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC the sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).

QY 1 CGTGTGCTGTGCTAGTCC 19
DB 213527 CGTGTGCTGTGCTAGTCC 213545

RESULT 38

Query Match 79.0%; Score 15.8; DB 10; Length 214520;
Best Local Similarity 89.5%; Pred. No. 1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20
DB 237103 GTGTGCTGTGCTAGTCCC 237121

RESULT 39
ADES59507/c
ID ADE59507 standard; DNA; 419 BP.
XX AC ADE59507;
XX DT 29-JAN-2004 (first entry)
XX DE Rat gene AA891308, SEQ ID NO 5402.
XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX

ADP43517
ID ADP43517 standard; DNA; 347001 BP.
XX AC ADP43517;
XX DT 09-SEP-2004 (first entry)
XX DE Human MAD1-like 1 DNA #7.
XX ds; gene; human.
XX KW Homo sapiens.
XX OS US2004115650-A1.
XX PN 17-JUN-2004.
XX PD 12-DEC-2002; 2002US-00319908.
XX PF 12-DEC-2002; 2002US-00319908.
XX PR (ISIS-) ISIS PHARM INC.
XX PA Dobie KW, Jain R;
XX PI WPI; 2004-449387/42.
XX DR New oligonucleotide compound that inhibits expression of MAD1-like 1,
XX PT useful for preparing a composition for treating hyperproliferative
XX PT disorder, e.g., cancer.
XX PS Example 15; SEQ ID NO 16; 206pp; English.
XX CC The invention relates to a new compound targeted to a nucleic acid
XX CC encoding MAD1-like 1 which specifically hybridises with the nucleic acid
XX CC encoding MAD1-like 1 and inhibits expression of MAD1-like 1. The
XX CC oligonucleotide compound is useful for preparing a composition for
XX CC treating hyperproliferative disorder, e.g. cancer. The present sequence
XX CC represents human MAD1-like 1 DNA.
XX SQ Sequence 347001 BP; 64746 A; 89476 C; 97054 G; 85725 T; 0 U; 10000 Other;

Query Match 79.0%; Score 15.8; DB 12; Length 347001;
Best Local Similarity 89.5%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGTGCTGTGCTAGTCCC 20
DB 237103 GTGTGCTGTGCTAGTCCC 237121

RESULT 39
ADES59507/c
ID ADE59507 standard; DNA; 419 BP.
XX AC ADE59507;
XX DT 29-JAN-2004 (first entry)
XX DE Rat gene AA891308, SEQ ID NO 5402.
XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX

```

PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; AA891308.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat DNA (shown in Table 2 of the
XX specification) which is differentially expressed during pain. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 419 BP; 113 A; 107 C; 87 G; 112 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 10; Length 419;
Best Local Similarity 94.1%; Pred. No. 7.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGCTGCTGCTAGTC 18
DB 175 GTGTGCTGCTGCTGTC 159

RESULT 40
ADJ10940/c
ID ADJ10940 standard; DNA; 448 BP.
XX
XX AC ADJ10940;
XX
XX 17-JUN-2004 (first entry)
XX
XX DE Recombinant tomato DNA to generate disease resistant plants SeqID 336.
XX
XX KW genetically engineered; plant; Avr-Pto; avirulent Pto;
XX bacterial speck disease; tomato; transgenic; biotic; abiotic stress; ds;
XX .Prf.
XX
XX OS Lycopersicon esculentum.
XX
XX PN US2004006787-A1.

```

```

XX 08-JAN-2004.
XX
XX 14-JAN-2003; 2003US-00341961.
XX
XX 14-JAN-2002; 2002US-0348792P.
XX 20-JUN-2002; 2002US-0390249P.
XX
XX (MART/) MARTIN G B.
XX (MYSO/) MYSOORE K K.
XX (CRAS/) CRASTA O R.
XX (FOLK/) FOLKERTS O.
XX (SWIR/) SWIRSKY P.
XX
XX Martin GB, Mysore KK, Crasta OR, Folkerts O, Swirsky P;
XX WPI; 2004-081759/08.
XX
XX New genetically engineered plant comprising a recombinant polynucleotide
XX showing expression associated with Avr-Pto mediated defense response,
XX useful in increasing resistance of plant against bacterial speck disease.
XX
XX Claim 5; SEQ ID NO 336; 176pp; English.
XX
XX This invention relates to a method of generating novel genetically
XX engineered plants. Specifically, it refers to compositions useful for
XX transforming plants with a recombinant polynucleotide or plant gene that
XX shows a specific pattern of expression associated with the Avr-Pto
XX mediated defence response. Avr-Pto is described as the avirulent Pto gene
XX found in Pseudomonas syringae pathovar tomato strain [strain Tl(A)],
XX which mediates bacterial speck disease in tomato plants. The present
XX invention describes identifying genes that are up- or down- regulated in
XX the defence response and that are involved in the interaction of Avr-Pto
XX with Pto and/ or Prf. As such, these novel recombinant polynucleotides
XX can be used to generate transgenic plants that are resistant to bacterial
XX speck disease and furthermore exhibit increased resistance against biotic
XX and abiotic stresses. This polynucleotide sequence is a recombinant
XX tomato polynucleotide used to generate genetically engineered, disease
XX resistant plants of the invention.
XX
XX Sequence 448 BP; 131 A; 102 C; 99 G; 115 T; 0 U; 1 Other;

Query Match 77.0%; Score 15.4; DB 12; Length 448;
Best Local Similarity 94.1%; Pred. No. 7.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGTGTGCTGCTGCTAGT 17
DB 116 CGTGTGCTGCTGCTAGT 100

RESULT 41
ABZ12515
ID ABZ12515 standard; DNA; 1644 BP.
XX
XX AC ABZ12515;
XX
XX 21-JAN-2003 (first entry)
XX
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 320.
XX
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
XX OS Arabidopsis thaliana.
XX
XX PN WO200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
XX 24-AUG-2000; 2000US-0227866P.
XX 26-JAN-2001; 2001US-0264647P.

```

PR 22-JUN-2001; 2001US-0300111P.
XX (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
PI WPI; 2002-304127/34.
XX
XX
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
XX Claim 144; SEQ ID NO 320; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
XX Sequence 1644 BP; 361 A; 358 C; 463 G; 462 T; 0 U; 0 Other;
SQ
Query Match 77.0%; Score 15.4; DB 6; Length 1644;
Best Local Similarity 94.1%; Pred. No. 8.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GTGTCTGTGCTAGTCCC 20
DB 1504 GTGTATGTGCTAGTCCC 1520
RESULT 42
ABL41879/c
ID ABL41879 standard; DNA; 1690 BP.
XX
XX ABL41879;
XX
XX 11-JUN-2002 (first entry)
DT
XX Nucleotide sequence of a linear polyol transporter.
DE
XX Linear polyol; mannitol; polyol transporter; carbon source; plant;
KW pathogen resistance; salt stress; sorbitol; dulcitol; galactitol;
KW inositol; ribitol; xylitol; ss.
XX
XX Arabidopsis thaliana.
OS
XX WO200204647-A1.
PN
XX 17-JAN-2002.
PD
XX 22-JUN-2001; 2001WO-FR001979.
PF
XX 11-JUL-2000; 2000FR-00009032.
PR
XX (CNRS) CENT NAT RECH SCI.
PA
XX Lemoine RRP, Noiraud NEJ;
PI WPI; 2002-154933/20.
XX
XX New polyol transporter protein from plants, for selecting transformed
PT cells and for imparting pathogen and salt-stress resistance to plants.
XX
XX Disclosure; Page 45; 66pp; French.
XX

CC The present sequence encodes a polypeptide which is a transporter of a
CC linear polyol. DNA encoding linear polyols is used to produce a selection
CC system for transformed cells, based on the use of polyols as the only
CC carbon source. It is also used to produce transgenic plants with
CC increased resistance to pathogens and salt stress. The use of a linear
CC polyol for selection eliminates the need for toxic selection reagents
CC such as antibiotics. Also, the linear polyol is not essential for the
CC plant, once selection has been made. The linear polyol has a main chain
CC of 5-8, preferably 6, carbon atoms and is selected from sorbitol,
CC dulcitol, galactitol, (myo-)inositol, ribitol, or xylitol
XX
XX Sequence 1690 BP; 479 A; 480 C; 365 G; 366 T; 0 U; 0 Other;
SQ
Query Match 77.0%; Score 15.4; DB 6; Length 1690;
Best Local Similarity 94.1%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GTGTCTGTGCTAGTCCC 20
DB 141 GTGTATGTGCTAGTCCC 125
RESULT 43
AAX20533
ID AAX20533 standard; DNA; 5819 BP.
XX
XX AAX20533;
AC
XX 05-MAY-1999 (first entry)
DT
XX Polynucleotide sequence from the genome of Treponema pallidum.
DE
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
KW
XX Treponema pallidum.
OS
XX WO9859034-A2.
PN
XX 30-DEC-1998.
PD
XX 23-JUN-1998; 98WO-US013041.
PF
XX 24-JUN-1997; 97US-0050667P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Fraser CM;
PI WPI; 1999-081273/07.
DR
XX New isolated Treponema pallidum nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
PT T. pallidum infections, particularly syphilis.
XX
XX Claim 1; Page 371-374; 1150pp; English.
XX
XX AAX20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for T. pallidum infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to Borrelia infections in animals, and for the production of
CC biosynthetic products such as enzymes
XX
XX Sequence 5819 BP; 1257 A; 1273 C; 1734 G; 1549 T; 0 U; 6 Other;
SQ
Query Match 77.0%; Score 15.4; DB 2; Length 5819;
Best Local Similarity 88.9%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GTGTCTGTGCTAGTCCC 19
DB 4816 GTGTGTCTGTGCTNCTCC 4833

RESULT 44

ADA02531/c
ID ADA02531 standard; DNA; 21981 BP.

XX AC

XX ADA02531;

XX DT

XX 06-NOV-2003 (first entry)

XX DE

XX Mouse Sox4 carcinoma associated gene, SEQ ID NO:1049.

XX KW

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;

XX KW

XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;

XX KW

XX gene; ds.

XX OS

XX Mus sp.

XX PN

XX WO2003057146-A2.

XX XX

XX 17-JUL-2003.

XX PD

XX 26-DEC-2002; 2002WO-US041414.

XX PF

XX 26-DEC-2001; 2001US-00035832.

XX PR

XX (SAGR-) SAGRES DISCOVERY.

XX PA

XX Morris DW;

XX PI

XX WPI; 2003-587068/55.

XX DR

XX New recombinant nucleic acid encoding carcinoma associated protein,

XX PT

XX useful for preparing compositions for treating carcinomas.

XX XX

XX Claim 1; SEQ ID NO 1049; 245pp; English.

XX CC

The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed murine CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ

XX Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;

XX Query Match

XX Best Local Similarity 77.0%; Score 15.4; DB 9; Length 21981;

XX Matches

XX 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX QY

XX 1 CGTGTCTGTGCTACT 17

XX Db

XX 8879 CGTGTCTGTGCTACT 8863

RESULT 45

ADB72269/c

ID ADB72269 standard; DNA; 21981 BP.

XX AC

XX ADB72269;

XX DT

XX 04-DEC-2003 (first entry)

XX XX

XX Mouse Sox4 gene.

XX DE

XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas; cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX OS

XX Mus sp.

XX XX

XX WO2003008583-A2.

XX PN

XX 30-JAN-2003.

XX PD

XX 26-DEC-2001; 2001WO-US051291.

XX PF

XX 02-MAR-2001; 2001US-00798586.

XX PR

XX 23-OCT-2001; 2001US-00004113.

XX PR

XX 08-NOV-2001; 2001US-00052482.

XX PR

XX 30-NOV-2001; 2001US-00997722.

XX PR

XX 20-DEC-2001; 2001US-00034650.

XX XX

XX (SAGR-) SAGRES DISCOVERY.

XX PA

XX Morris DW, Engelhard EK;

XX PI

XX WPI; 2003-239337/23.

XX DR

XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

XX PT

XX cancers, neoplasm, adenocarcinoma, or sarcomas.

XX XX

XX Claim 1; SEQ ID NO 97; 2304pp; English.

XX CC

The invention relates to a novel recombinant nucleic acid comprising a nucleotide sequence selected from any of the 660 sequences fully defined in the specification. A polynucleotide of the invention has cytostatic activity, and may have a use in gene therapy, or in a vaccine. The recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and sarcomas. The present sequence represents a mouse gene of the invention.

XX SQ

XX Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;

XX Query Match

XX Best Local Similarity 77.0%; Score 15.4; DB 10; Length 21981;

XX Matches

XX 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX QY

XX 1 CGTGTCTGTGCTACT 17

XX Db

XX 8879 CGTGTCTGTGCTACT 8863

RESULT 46

ADE95779/c

ID ADE95779 standard; DNA; 21981 BP.

XX XX

XX ADE95779;

XX AC

XX 12-FEB-2004 (first entry)

XX DT

XX Mouse Sox4 gene genomic DNA sequence.

XX DE

cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy; lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine; Sox4.

XX KW

XX OS

XX Mus sp.

XX XX

XX WO2003039484-A2.

XX PN

XX XX

PD 15-MAY-2003.
XX
PF 08-NOV-2002; 2002WO-US036071.
XX
XX 08-NOV-2001; 2001US-00052482.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
PI WPI; 2003-441462/41.
XX
DR New carcinoma associated nucleic acids and proteins, useful for screening
PT drug candidates, or for diagnosing and treating carcinomas, e.g.
PT lymphoma, breast cancer, prostate cancer or leukemia.
XX
XX Claim 1; SEQ ID NO 37; 793bp; English.
XX
XX This invention relates to novel recombinant nucleic acids for use in
CC diagnosis and treatment of cancer, especially carcinomas, as well as the
CC use of compositions in screening methods. The compositions of the
CC invention may have cytostatic activity whilst the disclosed sequences may
CC be useful for gene therapy. The carcinoma associated nucleic acids and
CC proteins are useful for diagnosing and treating carcinomas, for example
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC drug candidates or bioactive agents capable of binding to, or modulating
CC the activity of, a carcinoma associated protein. The present sequence is
CC the genomic DNA sequence of the mouse Sox4 gene which is a carcinoma
CC associated gene of the invention.
XX
XX Sequence 21981 BP; 5568 A; 4945 C; 5268 G; 5590 T; 0 U; 610 Other;
SQ
Query Match 77.0%; Score 15.4; DB 10; Length 21981;
Best Local Similarity 94.1%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGTCTACT 17
DB 8879 CGTGTCTGTGTCTACT 8863
RESULT 47
ABD32841
ID ABD32841 standard; DNA; 227448 BP.
XX
XX ABD32841;
AC
XX 18-NOV-2004 (first entry)
DT
XX Mouse cancer-associated genomic DNA MD17-014.
DE
XX Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX
XX Mus musculus.
OS
XX WO2004074320-A2.
PN
XX 02-SEP-2004.
PD
XX 17-FEB-2004; 2004WO-US004730.
XX
XX 14-FEB-2003; 2003US-00367094.
PR
XX 14-MAR-2003; 2003US-00388838.
PR
XX 15-APR-2003; 2003US-00417375.
PR
XX 13-JUN-2003; 2003US-00461862.
PR
XX 15-SEP-2003; 2003US-00663431.
PR
XX 15-DEC-2003; 2003US-00737318.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
PA
XX Morris DW, Morris DW, Malandro MS;
PI
XX

DR WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
XX disclosure; seqid 503; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 227448 BP; 54780 A; 54263 C; 56236 G; 60311 T; 0 U; 1858 Other;
SQ
Query Match 77.0%; Score 15.4; DB 13; Length 227448;
Best Local Similarity 94.1%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TGTGTCTGTGTCTACTCC 19
DB 198747 TGTGTCTGTGTCTACTCC 198763
RESULT 48
ABQ74179
ID ABQ74179 standard; DNA; 229354 BP.
XX
XX ABQ74179;
AC
XX 13-OCT-2002 (first entry)
DT
XX Human cytomegalovirus strain AD169 genomic sequence SEQ ID NO:64.
DE
XX Human cytomegalovirus; HCMV; CMV; Yeast artificial chromosome; YAC;
KW vaccine; immunisation; infection; antiinflammatory; immunosuppressive;
KW virucide; hepatotropic; congenital infection; jaundice; pneumonitis;
KW respiratory distress; convulsive seizure; mental retardation; hepatitis;
KW neurologic disability; asymptomatic adult infection; mononucleosis;
KW retinitis; acquired immune deficiency syndrome; AIDS; genome; ds.
XX
XX Human cytomegalovirus.
OS
XX
XX Key Location/Qualifiers
FH promoter 80996.81441
FT

FT FT /*tag= a
FT /note= "UL54 promoter"
FT 128318. 128750
FT /*tag= b
FT /note= "UL86 promoter"
XX WO200257437-A2.
PN 25-JUL-2002.
XX 30-OCT-2001; 2001WO-US047943.
XX 03-NOV-2000; 2000US-00705400.
PR (SCRI) SCRIPPS RES INST.
XX Ghazal P, Huang H;
XX WPI; 2002-590735/63.
DR Cytomegalovirus-yeast artificial chromosome (CMV-YAC) DNA molecule useful
XX as a vaccine for treating CMV infections e.g. jaundice, respiratory
PT distress, convulsive seizures, mononucleosis, hepatitis, pneumonitis and
PT retinitis.
XX Claim 5; Page 64-131; 132pp; English.
PS The present invention describes an isolated recombinant DNA molecule
XX comprising a yeast artificial chromosome (YAC) including at least a
CC portion of a human cytomegalovirus (HCMV) genome. Also described is a
CC composition suitable for use as a vaccine comprising the isolated
CC recombinant DNA molecule and an excipient. The YAC has antiinflammatory,
CC immunosuppressive, virucide, and hepatotropic activities, and can be used
CC in vaccine production. The YAC vector and the recombinant DNA molecule
CC are useful as vaccines for treating infections caused by CMV, e.g.
CC congenital infections such as jaundice, respiratory distress and
CC convulsive seizures, which may result in mental retardation, neurologic
CC disability or death; or asymptomatic adult infections such as
CC mononucleosis, hepatitis, pneumonitis and retinitis. The vaccine can also
CC be used for immuno-compromised patients suffering from acquired immune
CC deficiency syndrome (AIDS), or undergoing chemotherapy or tissue
CC rejection therapy after organ transplantation. The vector is also useful
CC as a tool for studying the replication of CMV using cell-line models and
CC the function of essential cis acting genes or nucleic acid sequences. The
CC YAC vector provides an ease and rapidity in introducing single or
CC multiple mutations into a CMV DNA sequence. The capacity of the vectors
CC to generate multiple mutations exceeds the capacity of other vectors,
CC e.g. bacterial artificial chromosome. The present sequence represents
CC human cytomegalovirus strain AD169 genomic sequence (also see Genbank
CC X17403.1), given in the present invention
XX Sequence 229354 BP; 49475 A; 64911 C; 66192 G; 48776 T; 0 U; 0 Other;
SQ Query Match 77.0%; Score 15.4; DB 6; Length 229354;
Best Local Similarity 94.1%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GGTGTCTGTGCTAGTC 18
DB 225853 GGTGTCTGTGCTAGTC 225869
RESULT 49
ACI35409/c
ID ACI35409 standard; DNA; 25 BP.
XX ACI35409;
XX ACI35409;
XX 13-OCT-2003 (first entry)
XX Human microarray DNA oligonucleotide SEQ ID NO 35400.
DE EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW

KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX Homo sapiens.
OS US2003104410-A1.
XX 05-JUN-2003.
XX 15-MAR-2002; 2002US-00098263.
XX 16-MAR-2001; 2001US-0276759P.
PR (AFFY-) AFFYMETRIX INC.
XX Mittmann MP;
PI WPI; 2003-567953/53.
DR New array of nucleic acid probes, useful for in situ hybridization, in
XX Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
PT Claim 1; SEQ ID NO 35400; 9pp; English.
XX The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html
XX Sequence 25 BP; 9 A; 5 C; 7 G; 4 T; 0 U; 0 Other;
SQ Query Match 76.0%; Score 15.2; DB 9; Length 25;
Best Local Similarity 85.0%; Pred. No. 6.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
DB 20 CATGTTCTGTCTAGTCCC 1
RESULT 50
ABN39874/c
ID ABN39874 standard; DNA; 60 BP.
XX ABN39874;
XX AC ABN39874;
XX 15-JUL-2002 (first entry)
XX Human spliced transcript detection oligonucleotide SEQ ID NO:12622.
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX

OS Homo sapiens.
XX
FN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB001903.
XX
XX 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
PS Example 1; SEQ ID NO 12622; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 BP; 18 A; 9 C; 20 G; 13 T; 0 U; 0 Other;

Query Match 76.0%; Score 15.2; DB 6; Length 60;
Best Local Similarity 85.0%; Pred.No. 7.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCC 20
DB 53 CCTGTGCTGTGCTAGTCC 34

Search completed: March 15, 2005, 18:32:45
Job time : 458 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2005, 15:23:29 ; Search time 1672 Seconds
(without alignments)
579.608 Million cell updates/sec

Title: US-09-980-953-256
Perfect score: 20
Sequence: 1 cgtgtctgtctgtatgcc 20
Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	124	6	ARI46422 Sequence
C 2	20	100.0	124	6	AR381504 Sequence
C 3	20	100.0	430	6	AX884562 Sequence
C 4	20	100.0	430	6	BD024172 Sequence
C 5	20	100.0	478	9	HSB7281
C 6	20	100.0	1112	6	ARS34819 Sequence
C 7	20	100.0	1112	9	HUMB72A
C 8	20	100.0	1120	6	AR030780 Sequence
C 9	20	100.0	1120	6	AR112747 Sequence
C 10	20	100.0	1120	6	AR146413 Sequence
C 11	20	100.0	1120	6	BD272169 Use of so
C 12	20	100.0	1120	6	AR196804 Sequence
C 13	20	100.0	1120	6	AR287728 Sequence
C 14	20	100.0	1120	6	AR374141 Sequence
C 15	20	100.0	1120	6	AR381495 Sequence
C 16	20	100.0	1120	6	AR527666 Sequence
C 17	20	100.0	1120	6	AX047043 Sequence
C 18	20	100.0	2781	9	BC040261 Homo sapi
C 19	20	100.0	164161	9	AC068630 Homo sapi

93	16.4	82.0	181585	9	AL139350	AL139350 Human DNA	166	15.8	79.0	1558	3	AY089564	AY089564 Drosophil
94	16.4	82.0	198267	2	AC137230	Rattus no	167	15.8	79.0	2269	9	BC017299	BC017299 Homo sapi
95	16.4	82.0	198602	2	AC114248	Rattus no	168	15.8	79.0	2269	9	AY026763	AY026763 Homo sapi
96	16.4	82.0	208595	2	AC139990	Rattus no	c 169	15.8	79.0	2284	5	BC072060	BC072060 Xenopus l
97	16.4	82.0	206741	10	AL732547	AL732547 Mouse DNA	c 170	15.8	79.0	2496	5	XLPOU3	X64835 X.laevis xl
98	16.4	82.0	213684	10	AC132616	AC132616 Mus muscu	c 171	15.8	79.0	2526	5	BC041298	BC041298 Xenopus l
99	16.4	82.0	213903	2	AC111731	AC111731 Rattus no	c 172	15.8	79.0	2927	5	BC080492	BC080492 Xenopus t
100	16.4	82.0	215530	2	AC105874	AC105874 Rattus no	c 173	15.8	79.0	3181	10	MMAJ740	AJ000740 Mus Muscu
101	16.4	82.0	223384	2	AC137440	AC137440 Rattus no	c 174	15.8	79.0	3186	9	AK123944	AK123944 Homo sapi
102	16.4	82.0	223693	2	AC103053	AC103053 Rattus no	175	15.8	79.0	3364	10	BC007130	BC007130 Mus muscu
103	16.4	82.0	225014	2	AC096045	AC096045 Rattus no	176	15.8	79.0	3545	6	CQ583988	CQ583988 Sequence
104	16.4	82.0	226604	2	AC094308	AC094308 Rattus no	c 177	15.8	79.0	4122	5	XLPOU3B	X96423 X.laevis PO
105	16.4	82.0	232843	2	AC133578	AC133578 Mus muscu	c 178	15.8	79.0	4282	10	MMAC13A	MM640 Mouse mRNA
106	16.4	82.0	243502	10	AL772402	AL772402 Mouse DNA	179	15.8	79.0	4409	3	AY157180	AY157180 Caryophyl
107	16.4	82.0	253401	2	AC097127	AC097127 Rattus no	c 180	15.8	79.0	5303	9	AC079143	AC079143 Homo sapi
108	16.4	82.0	266634	2	AC123487	AC123487 Rattus no	181	15.8	79.0	5856	6	CQ583994	CQ583994 Sequence
109	16.4	82.0	310003	9	AF178030	AF178030 Homo sapi	182	15.8	79.0	6413	3	AF047351	AF047351 Leishmani
110	16.4	82.0	347286	2	AC102740	AC102740 Mus muscu	183	15.8	79.0	15921	3	AF034856	AF034856 Drosophil
111	16	80.0	368	6	AX893238	AX893238 Sequence	184	15.8	79.0	24337	2	AC013130	AC013130 Drosophil
112	16	80.0	368	6	BD028771	BD028771 Sequence	185	15.8	79.0	26467	9	AP001472	AP001472 Homo sapi
113	16	80.0	50865	9	AL929444	AL929444 Human DNA	c 186	15.8	79.0	33931	9	AC022156	AC022156 Homo sapi
114	16	80.0	69206	2	AC020210	AC020210 Drosophil	187	15.8	79.0	36885	2	AC011015	AC011015 Leishmani
115	16	80.0	80449	2	AC114714	AC114714 Rattus no	188	15.8	79.0	36881	2	AC014041	AC014041 Drosophil
116	16	80.0	84350	2	AC016861	AC016861 Homo sapi	189	15.8	79.0	38817	10	AF427516	AF427516 Mus muscu
117	16	80.0	113810	9	AC068292	AC068292 Homo sapi	c 190	15.8	79.0	39855	9	AC020954	AC020954 Homo sapi
118	16	80.0	164689	2	AL953873	AL953873 Homo sapi	c 191	15.8	79.0	46064	5	AF461063	AF461063 Takifugu
119	16	80.0	168315	10	AC124451	AC124451 Mus muscu	c 192	15.8	79.0	48554	9	AC133104	AC133104 Homo sapi
120	16	80.0	168460	2	AC011585	AC011585 Homo sapi	193	15.8	79.0	58569	2	AC015033	AC015033 Drosophil
121	16	80.0	174274	2	BX957252	BX957252 Danilo rer	194	15.8	79.0	59351	2	AC100656	AC100656 Mus muscu
122	16	80.0	174274	2	BX957252	BX957252 Danilo rer	195	15.8	79.0	61170	2	AC068163	AC068163 Homo sapi
123	16	80.0	175436	2	AC122986	AC122986 Rattus no	c 196	15.8	79.0	61170	2	AC068163	AC068163 Homo sapi
124	16	80.0	175438	10	AC116598	AC116598 Mus muscu	197	15.8	79.0	67307	2	AC120598	Continuation (4 of
125	16	80.0	176195	3	AC012165	AC012165 Drosophil	c 198	15.8	79.0	69861	2	AC032033	Continuation (4 of
126	16	80.0	177911	2	AC021362	AC021362 Homo sapi	c 199	15.8	79.0	78456	10	AL627393	AC032033 Homo sapi
127	16	80.0	178810	9	AC106053	AC106053 Homo sapi	c 200	15.8	79.0	82827	9	AC079395	AL627393 Mouse DNA
128	16	80.0	180047	2	AC105174	AC105174 Mus muscu	201	15.8	79.0	82993	6	AX704274	AX704274 Sequence
129	16	80.0	182912	9	AC068060	AC068060 Homo sapi	202	15.8	79.0	91231	9	AL157364	AL157364 Human DNA
130	16	80.0	182913	9	AC068556	AC068556 Homo sapi	203	15.8	79.0	91489	9	AC010444	AC010444 Homo sapi
131	16	80.0	201862	2	AC150069	AC150069 Gallus ga	204	15.8	79.0	91825	2	AC019854	AC019854 Drosophil
132	16	80.0	203086	10	AC099948	AC099948 Mus muscu	c 205	15.8	79.0	97017	10	AL845423	AL845423 Mouse DNA
133	16	80.0	207432	3	AE003513	AE003513 Drosophil	c 206	15.8	79.0	98056	9	AC006061	AC006061 Homo sapi
134	16	80.0	220458	9	AC010320	AC010320 Homo sapi	207	15.8	79.0	110000	1	AE000516	Continuation (22 o
135	16	80.0	223708	2	AC102724	AC102724 Mus muscu	c 208	15.8	79.0	110000	2	AC008576	Continuation (2 of
136	16	80.0	239456	2	AC108995	AC108995 Rattus no	c 209	15.8	79.0	110000	2	AC079517	Continuation (3 of
137	16	80.0	243310	2	AC096208	AC096208 Rattus no	c 210	15.8	79.0	110000	2	AC132794	Continuation (4 of
138	16	80.0	245082	2	AC097237	AC097237 Rattus no	211	15.8	79.0	110000	3	AE003426	AE003426 Drosophil
139	16	80.0	249310	2	AC096891	AC096891 Rattus no	212	15.8	79.0	110000	8	CR382129	Continuation (9 of
140	16	80.0	257862	2	AC098100	AC098100 Rattus no	c 213	15.8	79.0	111061	9	AC078942	Continuation (9 of
141	16	80.0	258988	2	AC103322	AC103322 Rattus no	214	15.8	79.0	114000	2	AC125412	AC078942 Homo sapi
142	16	80.0	270456	2	AC111575	AC111575 Rattus no	c 215	15.8	79.0	116908	2	AC018365	AC125412 Leishmani
143	15.8	79.0	328	11	BV074663	BV074663 S208P6019	c 216	15.8	79.0	123244	10	BX004791	AC018365 Mus muscu
144	15.8	79.0	336	11	MMST5162	236700 M.musculus	217	15.8	79.0	127883	10	AC114434	BX004791 Mouse DNA
145	15.8	79.0	430	4	BTMSAT76	D71569 B.taurus mi	218	15.8	79.0	128614	10	BX510300	AC114434 Rattus no
146	15.8	79.0	442	3	OCSTINEF	D32088 Octopus vul	c 219	15.8	79.0	131000	9	AF001017	AP001017 Homo sapi
147	15.8	79.0	458	4	BTMSAT72	D71565 B.taurus mi	220	15.8	79.0	131264	2	AC104198	AL034198 Mus muscu
148	15.8	79.0	460	3	OCSTINEL	D32094 Octopus vul	221	15.8	79.0	131819	9	HS529N6	AL034376 Human DNA
149	15.8	79.0	471	9	HSTCB4	X74844 H.sapiens (c 222	15.8	79.0	132176	2	AC150122	AC150122 Gallus ga
150	15.8	79.0	609	9	HS3339102	AJ339102 Homo sapi	c 223	15.8	79.0	135276	2	AC073603	AC150122 Gallus ga
151	15.8	79.0	628	11	BV021838	BV021838 S212P6375	224	15.8	79.0	136010	2	AC150128	AC150128 Gallus ga
152	15.8	79.0	642	9	HS3339311	AJ339311 Homo sapi	c 225	15.8	79.0	137286	4	AL773527	AL773527 Pig DNA s
153	15.8	79.0	642	9	HS3339313	AJ339313 Homo sapi	c 226	15.8	79.0	137898	2	AC140552	AC140552 Mus muscu
154	15.8	79.0	656	9	HS342373	AJ342373 Homo sapi	c 227	15.8	79.0	138048	2	AC113575	AL13575 Tetraodon
155	15.8	79.0	662	9	HS332969	AJ322969 Homo sapi	c 228	15.8	79.0	138217	10	AL669905	AL669905 Mouse DNA
156	15.8	79.0	665	9	HS342510	AJ342510 Homo sapi	c 229	15.8	79.0	138320	2	AC150081	AC150081 Gallus ga
157	15.8	79.0	667	9	HS3329574	AJ329574 Homo sapi	c 230	15.8	79.0	139894	2	AC151208	AC151208 Bos tauru
158	15.8	79.0	678	9	HS340234	AJ340234 Homo sapi	c 231	15.8	79.0	141073	9	AC105108	AC105108 Homo sapi
159	15.8	79.0	682	9	HS340233	AJ340233 Homo sapi	c 232	15.8	79.0	141766	2	AC148652	AC148652 Medicago
160	15.8	79.0	729	9	HS330600	AJ330600 Homo sapi	233	15.8	79.0	143895	2	CR407557	CR407557 Danilo rer
161	15.8	79.0	738	9	HS3338205	AJ338205 Homo sapi	c 234	15.8	79.0	144699	2	AC125881	AC125881 Rattus no
162	15.8	79.0	765	5	CR391089	CR391089 Gallus ga	c 235	15.8	79.0	145709	2	AC069288	AC069288 Homo sapi
163	15.8	79.0	931	3	OCSTINEA	D32083 Octopus vul	c 236	15.8	79.0	147634	9	AC055740	AC055740 Homo sapi
164	15.8	79.0	1105	9	CR391089	BC063580 Homo sapi	c 237	15.8	79.0	149427	2	AC121385	AC121385 Rattus no
165	15.8	79.0	1217	9	AY177407	AY177407 Homo sapi	c 238	15.8	79.0	150069	2	AC041042	AC041042 Homo sapi

c 239	15.8	79.0	151302	9	AC073617	AC073617 Homo sapi
c 240	15.8	79.0	152827	2	AC143347	AC143347 Homo sapi
c 241	15.8	79.0	153565	10	AC080016	AC080016 Mus muscu
c 242	15.8	79.0	154154	2	AC121679	AC121679 Rattus no
c 243	15.8	79.0	155001	2	AL356674	AL356674 Homo sapi
c 244	15.8	79.0	155670	3	AC012097	AC012097 Drosophill
c 245	15.8	79.0	155685	3	AC098575	AC098575 Drosophill
c 246	15.8	79.0	156684	10	AC058787	AC058787 Mus muscu
c 247	15.8	79.0	157091	2	AC027772	AC027772 Homo sapi
c 248	15.8	79.0	157435	5	AL935029	AL935029 Zebrafish
c 249	15.8	79.0	157633	9	AC008687	AC008687 Homo sapi
c 250	15.8	79.0	157659	2	AC143345	AC143345 Homo sapi
c 251	15.8	79.0	157733	2	AC055810	AC055810 Homo sapi
c 252	15.8	79.0	157812	9	AC087782	AC087782 Homo sapi
c 253	15.8	79.0	157921	2	AC027726	AC027726 Homo sapi
c 254	15.8	79.0	158121	2	AC147848	AC147848 Homo sapi
c 255	15.8	79.0	158377	2	AC147522	AC147522 Ocolemur
c 256	15.8	79.0	161354	2	AC113374	AC113374 Homo sapi
c 257	15.8	79.0	162835	10	AC121866	AC121866 Mus muscu
c 258	15.8	79.0	163724	2	CR788229	CR788229 Danio rer
c 259	15.8	79.0	163923	2	AC074278	AC074278 Homo sapi
c 260	15.8	79.0	164511	9	AC004894	AC004894 Homo sapi
c 261	15.8	79.0	164635	2	AC143346	AC143346 Homo sapi
c 262	15.8	79.0	165002	2	AC150642	AC150642 Bos tauru
c 263	15.8	79.0	165033	2	AC087125	AC087125 Mus muscu
c 264	15.8	79.0	165212	10	AL596125	AL596125 Mouse DNA
c 265	15.8	79.0	165378	2	AC143332	AC143332 Homo sapi
c 266	15.8	79.0	166065	2	AC149014	AC149014 Papio anu
c 267	15.8	79.0	166171	2	AL359550	AL359550 Homo sapi
c 268	15.8	79.0	166787	9	AC055723	AC055723 Homo sapi
c 269	15.8	79.0	167053	10	AC124447	AC124447 Mus muscu
c 270	15.8	79.0	167905	2	AC116606	AC116606 Homo sapi
c 271	15.8	79.0	168277	2	AC147934	AC147934 Ocolemur
c 272	15.8	79.0	168509	2	AC115344	AC115344 Rattus no
c 273	15.8	79.0	168671	9	AC019247	AC019247 Homo sapi
c 274	15.8	79.0	168779	9	AC026405	AC026405 Homo sapi
c 275	15.8	79.0	169046	5	AL928828	AL928828 Zebrafish
c 276	15.8	79.0	169155	5	BX088685	BX088685 Zebrafish
c 277	15.8	79.0	169457	2	AC009216	AC009216 Drosophill
c 278	15.8	79.0	170102	9	AC009237	AC009237 Homo sapi
c 279	15.8	79.0	170323	10	AC124459	AC124459 Mus muscu
c 280	15.8	79.0	171451	10	AC079444	AC079444 Mus muscu
c 281	15.8	79.0	171557	2	AC147276	AC147276 Pan trogl
c 282	15.8	79.0	171961	9	AC113408	AC113408 Homo sapi
c 283	15.8	79.0	171982	2	AC120916	AC120916 Rattus no
c 284	15.8	79.0	172253	2	CR847922	CR847922 Danio rer
c 285	15.8	79.0	172298	2	AC055121	AC055121 Homo sapi
c 286	15.8	79.0	172382	2	AC149046	AC149046 Papio anu
c 287	15.8	79.0	172963	9	AC100763	AC100763 Homo sapi
c 288	15.8	79.0	173126	9	AC005988	AC005988 Homo sapi
c 289	15.8	79.0	173667	2	AC134316	AC134316 Lemur cat
c 290	15.8	79.0	173882	5	BX640462	BX640462 Zebrafish
c 291	15.8	79.0	175231	10	AC034108	AC034108 Mus muscu
c 292	15.8	79.0	176384	2	AC150507	AC150507 Bos tauru
c 293	15.8	79.0	176767	9	AF196779	AF196779 Homo sapi
c 294	15.8	79.0	177956	2	AC110916	AC110916 Mus muscu
c 295	15.8	79.0	178208	2	AC136432	AC136432 Homo sapi
c 296	15.8	79.0	178346	2	AC134539	AC134539 Mus muscu
c 297	15.8	79.0	178443	10	AL592185	AL592185 Mouse DNA
c 298	15.8	79.0	178457	5	AL935031	AL935031 Zebrafish
c 299	15.8	79.0	178728	2	AC143348	AC143348 Homo sapi
c 300	15.8	79.0	179145	9	AC009785	AC009785 Homo sapi
c 301	15.8	79.0	179155	9	AL353748	AL353748 Human DNA
c 302	15.8	79.0	179192	2	AC135529	AC135529 Rattus no
c 303	15.8	79.0	179501	9	AL591543	AL591543 Human DNA
c 304	15.8	79.0	179592	2	AC009723	AC009723 Homo sapi
c 305	15.8	79.0	179649	3	AC010045	AC010045 Drosophill
c 306	15.8	79.0	179711	2	AC019312	AC019312 Homo sapi
c 307	15.8	79.0	179941	2	AC150586	AC150586 Bos tauru
c 308	15.8	79.0	180787	3	AC009375	AC009375 Drosophill
c 309	15.8	79.0	180914	9	CNS01051	AL121612 Human chr
c 310	15.8	79.0	181111	10	AC129301	AC129301 Mus muscu
c 311	15.8	79.0	181162	5	AC144487	AC144487 Gasterost
c 312	15.8	79.0	182553	2	AC149646	AC149646 Bos tauru
c 313	15.8	79.0	182661	10	AC121602	AC121602 Mus muscu
c 314	15.8	79.0	182829	10	AC135859	AC135859 Mus muscu
c 315	15.8	79.0	183239	2	AC015906	AC015906 Homo sapi
c 316	15.8	79.0	183273	9	AC093898	AC093898 Homo sapi
c 317	15.8	79.0	183586	9	AC092562	AC092562 Papio ham
c 318	15.8	79.0	183720	2	AC092912	AC092912 Homo sapi
c 319	15.8	79.0	183897	2	AC150665	AC150665 Bos tauru
c 320	15.8	79.0	184163	2	AC149723	AC149723 Bos tauru
c 321	15.8	79.0	184798	2	AC139981	AC139981 Rattus no
c 322	15.8	79.0	185800	2	AC135224	AC135224 Gallus ga
c 323	15.8	79.0	186058	2	AC112245	AC112245 Homo sapi
c 324	15.8	79.0	186144	2	AC069048	AC069048 Homo sapi
c 325	15.8	79.0	186258	5	BX640480	BX640480 Zebrafish
c 326	15.8	79.0	186347	2	AC107710	AC107710 Mus muscu
c 327	15.8	79.0	186748	3	AC012163	AC012163 Drosophill
c 328	15.8	79.0	187058	10	AC126410	AC126410 Mus muscu
c 329	15.8	79.0	187343	2	AC016688	AC016688 Homo sapi
c 330	15.8	79.0	187344	9	AL354873	AL354873 Human DNA
c 331	15.8	79.0	187450	10	BX322546	BX322546 Mouse DNA
c 332	15.8	79.0	188646	2	AC141652	AC141652 Rattus no
c 333	15.8	79.0	189023	9	AL592114	AL592114 Human DNA
c 334	15.8	79.0	189263	10	AC134595	AC134595 Mus muscu
c 335	15.8	79.0	189650	10	AC136091	AC136091 Rattus no
c 336	15.8	79.0	191674	4	CR536601	CR536601 Platypus
c 337	15.8	79.0	191923	9	AC109480	AC109480 Homo sapi
c 338	15.8	79.0	192015	2	AC119223	AC119223 Mus muscu
c 339	15.8	79.0	192556	2	AC103942	AC103942 Mus muscu
c 340	15.8	79.0	192569	2	AC133928	AC133928 Lemur cat
c 341	15.8	79.0	192687	2	AC129244	AC129244 Rattus no
c 342	15.8	79.0	192994	2	CR318588	CR318588 Danio rer
c 343	15.8	79.0	193147	2	AC149200	AC149200 Macaca mu
c 344	15.8	79.0	193286	2	AC110119	AC110119 Rattus no
c 345	15.8	79.0	193580	2	AC117284	AC117284 Rattus no
c 346	15.8	79.0	194109	10	AC121819	AC121819 Mus muscu
c 347	15.8	79.0	194171	2	AC126452	AC126452 Mus muscu
c 348	15.8	79.0	195728	2	CR384051	CR384051 Danio rer
c 349	15.8	79.0	197229	2	AC021162	AC021162 Homo sapi
c 350	15.8	79.0	197679	10	AC128663	AC128663 Mus muscu
c 351	15.8	79.0	198116	2	AC123155	AC123155 Rattus no
c 352	15.8	79.0	198919	10	AC140277	AC140277 Mus muscu
c 353	15.8	79.0	198996	2	AC134845	AC134845 Mus muscu
c 354	15.8	79.0	200946	2	AC148421	AC148421 Meleagris
c 355	15.8	79.0	200985	3	AC010060	AC010060 Drosophill
c 356	15.8	79.0	201485	2	AC134402	AC134402 Mus muscu
c 357	15.8	79.0	201602	2	AC074146	AC074146 Mus muscu
c 358	15.8	79.0	202083	2	AC023833	AC023833 Mus muscu
c 359	15.8	79.0	202086	10	AC080015	AC080015 Mus muscu
c 360	15.8	79.0	202151	10	AC102358	AC102358 Mus muscu
c 361	15.8	79.0	203495	2	BX936436	BX936436 Danio rer
c 362	15.8	79.0	203765	9	AC087289	AC087289 Homo sapi
c 363	15.8	79.0	203829	5	AC144485	AC144485 Gasterost
c 364	15.8	79.0	204214	10	AL672251	AL672251 Mouse DNA
c 365	15.8	79.0	204239	10	AC114819	AC114819 Mus muscu
c 366	15.8	79.0	204892	2	AC151244	AC151244 Bos tauru
c 367	15.8	79.0	205503	2	AC148001	AC148001 Mus muscu
c 368	15.8	79.0	205553	2	AC118516	AC118516 Rattus no
c 369	15.8	79.0	205985	2	AC145399	AC145399 Rattus no
c 370	15.8	79.0	207424	10	AC090489	AC090489 Genomic s
c 371	15.8	79.0	208214	10	AL645908	AL645908 Mouse DNA
c 372	15.8	79.0	209877	9	AC012509	AC012509 Homo sapi
c 373	15.8	79.0	210206	2	AC139598	AC139598 Rattus no
c 374	15.8	79.0	210206	2	AC139598	AC139598 Rattus no
c 375	15.8	79.0	211350	2	AC112604	AC112604 Rattus no
c 376	15.8	79.0	211763	10	AC123558	AC123558 Mus muscu
c 377	15.8	79.0	212120	2	AC023414	AC023414 Homo sapi
c 378	15.8	79.0	212628	2	AC097283	AC097283 Rattus no
c 379	15.8	79.0	213746	2	CR847568	CR847568 Danio rer
c 380	15.8	79.0	214527	9	CNS00007	AL049839 Human chr
c 381	15.8	79.0	215196	2	AC073796	AC073796 Mus muscu
c 382	15.8	79.0	215199	2	AC149052	AC149052 Mus muscu
c 383	15.8	79.0	215863	2	AC136812	AC136812 Rattus no
c 384	15.8	79.0	215914	2	AC127875	AC127875 Rattus no

C 385	15.8	79.0	216807	10	AC079681	Mus muscu
C 386	15.8	79.0	219121	2	AC134718	Rattus no
C 387	15.8	79.0	219409	2	AC134858	Mus muscu
C 388	15.8	79.0	219747	2	AC098304	Rattus no
C 389	15.8	79.0	220788	2	AC130904	Rattus no
C 390	15.8	79.0	220805	2	AC118515	Rattus no
C 391	15.8	79.0	220931	2	AC116066	Rattus no
C 392	15.8	79.0	220931	2	AC131405	Rattus no
C 393	15.8	79.0	223344	2	AC106193	Rattus no
C 394	15.8	79.0	223345	2	AC123324	Rattus no
C 395	15.8	79.0	223971	2	AC115753	Mus muscu
C 396	15.8	79.0	223979	2	AC096151	Rattus no
C 397	15.8	79.0	224592	2	AC074209	Mus muscu
C 398	15.8	79.0	224946	2	AC115413	Rattus no
C 399	15.8	79.0	225005	2	AC107607	Rattus no
C 400	15.8	79.0	225274	2	AC130251	Rattus no
C 401	15.8	79.0	225924	2	AC074210	Mus muscu
C 402	15.8	79.0	226640	2	AC097833	Rattus no
C 403	15.8	79.0	226865	2	AC098153	Rattus no
C 404	15.8	79.0	227364	2	AC113732	Rattus no
C 405	15.8	79.0	227833	2	AC128446	Rattus no
C 406	15.8	79.0	227834	4	AC150561	Bos tauri
C 407	15.8	79.0	228858	2	AC147751	Otolemur
C 408	15.8	79.0	229063	2	AC114701	Rattus no
C 409	15.8	79.0	229131	2	AC074149	Mus muscu
C 410	15.8	79.0	229482	2	AC074149	Mus muscu
C 411	15.8	79.0	230783	2	AC118913	Rattus no
C 412	15.8	79.0	231234	2	AC148222	Colobus g
C 413	15.8	79.0	231722	2	AC097596	Rattus no
C 414	15.8	79.0	231864	2	AC129642	Rattus no
C 415	15.8	79.0	232473	10	AC128896	Mus muscu
C 416	15.8	79.0	233566	2	AC118800	Rattus no
C 417	15.8	79.0	233784	2	AC130877	Rattus no
C 418	15.8	79.0	234399	2	AC117147	Rattus no
C 419	15.8	79.0	234655	2	AC096021	Rattus no
C 420	15.8	79.0	234681	2	AC095940	Rattus no
C 421	15.8	79.0	234842	2	AC127153	Rattus no
C 422	15.8	79.0	234967	2	AC134282	Rattus no
C 423	15.8	79.0	235025	2	AC114149	Rattus no
C 424	15.8	79.0	236261	2	AC115280	Rattus no
C 425	15.8	79.0	236456	2	AC036147	Mus muscu
C 426	15.8	79.0	237327	2	AC103454	Rattus no
C 427	15.8	79.0	238050	2	AC094265	Rattus no
C 428	15.8	79.0	239570	2	AC137983	Mus muscu
C 429	15.8	79.0	239946	2	AC103505	Rattus no
C 430	15.8	79.0	240142	2	AC134295	Rattus no
C 431	15.8	79.0	240596	2	AC095941	Rattus no
C 432	15.8	79.0	241069	2	AC124894	Rattus no
C 433	15.8	79.0	241721	2	AC111729	Rattus no
C 434	15.8	79.0	242465	2	AC128332	Rattus no
C 435	15.8	79.0	243354	5	BX470172	Zebrafish
C 436	15.8	79.0	243537	2	AC132629	Rattus no
C 437	15.8	79.0	243981	2	AC111281	Rattus no
C 438	15.8	79.0	244068	2	AC131345	Rattus no
C 439	15.8	79.0	244377	2	AC110661	Rattus no
C 440	15.8	79.0	245157	2	AC119557	Rattus no
C 441	15.8	79.0	245417	2	AC105891	Rattus no
C 442	15.8	79.0	245737	2	AC122654	Rattus no
C 443	15.8	79.0	246006	2	AC102141	Mus muscu
C 444	15.8	79.0	246772	2	AC112469	Rattus no
C 445	15.8	79.0	247046	2	AC119441	Rattus no
C 446	15.8	79.0	247146	2	AC134757	Rattus no
C 447	15.8	79.0	248196	2	AC127847	Rattus no
C 448	15.8	79.0	248294	2	AC097690	Rattus no
C 449	15.8	79.0	249364	2	AC098392	Rattus no
C 450	15.8	79.0	249549	2	AC109236	Mus muscu
C 451	15.8	79.0	249781	2	AC098954	Rattus no
C 452	15.8	79.0	250529	2	AC097607	Rattus no
C 453	15.8	79.0	251478	2	AC105849	Rattus no
C 454	15.8	79.0	251495	2	AC113722	Rattus no
C 455	15.8	79.0	251740	2	AC123460	Rattus no
C 456	15.8	79.0	252041	2	AC114432	Rattus no
C 457	15.8	79.0	252257	2	AC096047	Rattus no

C 458	15.8	79.0	252580	2	AC121227	Rattus no
C 459	15.8	79.0	252775	2	AC109711	Rattus no
C 460	15.8	79.0	253922	2	AC111987	Rattus no
C 461	15.8	79.0	254380	2	AC105841	Rattus no
C 462	15.8	79.0	254380	10	AC118234	Mus muscu
C 463	15.8	79.0	256220	2	AC109981	Rattus no
C 464	15.8	79.0	257333	2	AC132722	Rattus no
C 465	15.8	79.0	257396	2	AC130569	Rattus no
C 466	15.8	79.0	257645	2	AC103094	Rattus no
C 467	15.8	79.0	262581	2	AC116230	Rattus no
C 468	15.8	79.0	263939	2	AC095238	Rattus no
C 469	15.8	79.0	265067	2	AC097113	Rattus no
C 470	15.8	79.0	265515	2	AC110852	Rattus no
C 471	15.8	79.0	266762	2	AC109853	Rattus no
C 472	15.8	79.0	271034	2	AC102996	Rattus no
C 473	15.8	79.0	271496	2	AC142073	Rattus no
C 474	15.8	79.0	272535	5	EX005212	Zebrafish
C 475	15.8	79.0	275277	2	AC129438	Rattus no
C 476	15.8	79.0	277347	2	AC130968	Rattus no
C 477	15.8	79.0	281117	2	AC110645	Rattus no
C 478	15.8	79.0	289140	3	AE003533	Drosophil
C 479	15.8	79.0	290138	2	AC097202	Rattus no
C 480	15.8	79.0	291050	1	EX248340	Mycobacte
C 481	15.8	79.0	295466	2	AC123317	Rattus no
C 482	15.8	79.0	299804	3	AE003507	Drosophil
C 483	15.8	79.0	302050	3	AE003522	Drosophil
C 484	15.8	79.0	313499	2	AC119366	Rattus no
C 485	15.8	79.0	318535	2	AC107011	Rattus no
C 486	15.8	79.0	327876	2	AC095476	Rattus no
C 487	15.8	79.0	327876	2	AC095476	Rattus no
C 488	15.8	79.0	328863	2	AL627202	Homo sapi
C 489	15.8	79.0	336392	2	AC130441	Rattus no
C 490	15.8	79.0	340000	9	AP001759	Homo sapi
C 491	15.8	79.0	342336	1	EX842578	Mycobacte
C 492	15.8	79.0	346186	11	AU049694	Rattus no
C 493	15.4	77.0	334	11	G47212	Zi15893_1 Ze
C 494	15.4	77.0	351	5	AY096855	Glyphorin
C 495	15.4	77.0	351	5	AY096871	Glyphorin
C 496	15.4	77.0	351	5	AY096872	Glyphorin
C 497	15.4	77.0	351	5	AY096873	Glyphorin
C 498	15.4	77.0	351	5	AY096874	Glyphorin
C 499	15.4	77.0	351	5	AY096875	Glyphorin
C 500	15.4	77.0	351	5	AY096875	Glyphorin

ALIGNMENTS

RESULT 1	ARI146422/c	Sequence 39 from patent US 6218510.	124 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	ARI146422					
DEFINITION	ARI146422					
ACCESSION	ARI146422.1	GI:15109611				
VERSION						
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 124)					
AUTHORS	Sharpe,A.H., Borriello,F., Freeman,G.J. and Nadler,L.M.					
TITLE	B7-1 and B7-2 polypeptides					
JOURNAL	Patent: US 6218510-A 39 17-APR-2001;					
FEATURES	Location/Qualifiers					
source	1..124					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						
Query Match	100.0%;	Score 20;	DB 6;	Length 124;		
Best Local Similarity	100.0%;	Pred. No. 44;				
Matches	20;	Conservative	0;	Mismatches	0;	Gaps 0;
QY	1	CGTGTGTCGTGCTAGTCCC	20			

```

Db      57  CGTGTCTGTGCTAGTCCC 38
|||||
RESULT 2
AR381504/c
LOCUS      AR381504      124 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 39 from patent US 6608180.
ACCESSION  AR381504
VERSION     AR381504.1  GI:40089565
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 124)
AUTHORS     Sharpe,A.H., Borriello,F., Freeman,G.J. and Nadler,L.M.
TITLE       B7-specific antibodies
JOURNAL     Patent: US 6608180-A 39 19-AUG-2003;
FEATURES    Location/Qualifiers
             source
               1..124
               /organism="unknown"
               /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTCTGTGCTAGTCCC 20
|||||
Db      57  CGTGTCTGTGCTAGTCCC 38
|||||
RESULT 3
AX884562/c
LOCUS      AX884562      430 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 425 from Patent EP1033401.
ACCESSION  AX884562
VERSION     AX884562.1  GI:40039847
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE       Expressed sequence tags and encoded human proteins
JOURNAL     Patent: EP 1033401-A 425 06-SEP-2000;
FEATURES    Location/Qualifiers
             source
               1..430
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
               125..430
               /note="unnamed protein product"
               /codon_start=1
               /protein_id="CAE98676.1"
               /db_xref="GI:40039848"
               /translation="MDPQCTMGLSNILFWAFLLSGAAPLKIOAFNFTADLPQCFAN
               "
               SQQSLSLWVFWQDENLVLEVLGKFKDSVHSKYMGRTSFSDSDSWTLRLHLQI
             sig_peptide
               125..199
               /note="score 7.1 seq LFVMAFLLSGAAP/LK"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTCTGTGCTAGTCCC 20
|||||
Db      75  CGTGTCTGTGCTAGTCCC 56
|||||
RESULT 4
BD024172/c
LOCUS      BD024172      430 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION  BD024172
VERSION     BD024172.1  GI:22565395
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 430)
AUTHORS     Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE       Sequence tag and encoded human protein
JOURNAL     Patent: JP 2001269182-A 418 02-OCT-2001;
FEATURES    GENSET
             OS Homo sapiens (human)
             PN JP 2001269182-A/418
             PD 02-OCT-2001
             PF 24-FEB-2000 JP 2000118773
             PR 26-FEB-1999 US 60/122487
             PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCCLAIR,JEAN YVES
             PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
             PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
             G06F15/40
             CC score 7.1
             CC seq LFVMAFLLSGAAP/LK
             FH Key Location/Qualifiers
             FT CDS 125..430
             FT sig_peptide 125..199.
             Location/Qualifiers
               1..430
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTGTCTGTGCTAGTCCC 20
|||||
Db      75  CGTGTCTGTGCTAGTCCC 56
|||||
RESULT 5
HSB72S1/c
LOCUS      HSB72S1      478 bp      DNA      linear      PRI 02-FEB-1996
DEFINITION Human CTLA-4 counter-receptor B7.2 (B7.2) gene, exon 1.
ACCESSION  U17715
VERSION     U17715.1  GI:808025
KEYWORDS    1 of 8
SEGMENT     Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 478)
AUTHORS     Jellis,C.L., Wang,S.S., Rennert,P., Borriello,F., Sharpe,A.H.,
             Green,N.R. and Gray,G.S.
TITLE       Genomic organization of the gene coding for the costimulatory human
             B-lymphocyte antigen B7-2 (CD86)
JOURNAL     Immunogenetics 42 (2), 85-89 (1995)
MEDLINE     9531831
PUBMED      7541777
REFERENCE   2 (bases 173 to 292)
AUTHORS     Freeman,G.J., Gribben,J.G., Boussiotis,V.A., Ng,J.W., Restivo,V.A.
             Jr., Lombard,L.A., Gray,G.S. and Nadler,L.M.

```

TITLE Cloning of B7-2: a CTLA-4 counter-receptor that costimulates human T cell proliferation
JOURNAL Science 262 (5135), 909-911 (1993)
MEDLINE 94053735
PUBMED 7694363
REFERENCE 3 (sites)
AUTHORS Azuma,M., Ito,D., Yagita,H., Okumura,K., Phillips,J.H., Lanier,L.L. and Somoza,C.
TITLE B70 antigen is a second ligand for CTLA-4 and CD28
JOURNAL Nature 366 (6450), 76-79 (1993)
MEDLINE 94050123
PUBMED 7694153
REFERENCE 4 (bases 1 to 478)
AUTHORS Jellis,C.L.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-1994) Cindy L. Jellis, Molecular Biology, Repligen Corporation, One Kendall Square, Building 700, Cambridge, MA 02139, USA

FEATURES
Location/Qualifiers
source
1..478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="DMPC-HFF no.1-1390-A1 hB7.2 genomic DNA (exon 1)"
/sex="male"
/cell_type="fibroblast"
/tissue_type="foreskin"
/clone_lib="Du Pont Merck Pharmaceutical Company Human foreskin Fibroblast PI Library no.1"
173..292
/gene="B7.2"
/note="alternative 5' exon"
/citation=[2]
/citation=[1]
/number=1

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 478;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGTGTCTGTGCTAGTCCC 20
|||||
Db 229 CGTGTCTGTGCTAGTCCC 210

RESULT 6
LOCUS AR534819/c 1112 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 25 from patent US 6734172.
ACCESSION AR534819
VERSION AR534819.1 GI:53925465
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1112)
AUTHORS Scholler,N.B., Disis,M.L., Hellstrom,I. and Hellstrom,K.E.
TITLE Surface receptor antigen vaccines
JOURNAL Patent: US 6734172-A 25 11-MAY-2004;
FEATURES Location/Qualifiers
source
1..1112
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1112;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGTGTCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 7
LOCUS HUMB72A/c 1112 bp mRNA linear PRI 31-DEC-1994
DEFINITION Human CTLA4 counter-receptor (B7-2) mRNA, complete cds.
ACCESSION L25259
VERSION L25259.1 GI:416368
KEYWORDS CTLA4 counter-receptor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1112)
AUTHORS Freeman,G.J., Gribben,J.G., Boussiotis,V.A., Ng,J.W., Restivo,V.A. Jr., Lombard,L.A., Gray,G.S. and Nadler,L.M.
TITLE Cloning of B7-2: a CTLA-4 counter-receptor that costimulates human T cell proliferation
JOURNAL Science 262 (5135), 909-911 (1993)
MEDLINE 94053735
PUBMED 7694363
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES Location/Qualifiers
source
1..1112
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="B lymphocyte"
1..1112
/gene="B7-2"
107..1096
/gene="B7-2"
/codon_start=1
/product="CTLA4 counter-receptor"
/protein_id="AA58389.1"
/db_xref="GI:416369"
/translation="MDPQTMGLSNILFVMAFLSGLAPLKIQAYFNETADLPQCFAN
SONQSLSELVVFQDQENLVNLYLGRKFDVSHVSGRTSFDSDSWLRLNLQI
KDGLYQCIHHKPTQMIHOMNSLSLANFQPEIVISNITENVINLTCSS
HGYPEPKMSVLLRTKNSITIEYDGMOKSDONTVELYDVISLSVSPDVTNNMTIFC
ILETDKTRLLSSPFSIELEDPQPPDHPIHTAVLPTVILCVAVFCILIMKWKKKRP
RNSYKGTNTWREESQTKKREKIHIPERSDEAQRVFKSKTSKSDKSDTCF"

source

1..1112

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_type="B lymphocyte"

1..1112

/gene="B7-2"

107..1096

/gene="B7-2"

/codon_start=1

/product="CTLA4 counter-receptor"

/protein_id="AA58389.1"

/db_xref="GI:416369"

/translation="MDPQTMGLSNILFVMAFLSGLAPLKIQAYFNETADLPQCFAN
SONQSLSELVVFQDQENLVNLYLGRKFDVSHVSGRTSFDSDSWLRLNLQI
KDGLYQCIHHKPTQMIHOMNSLSLANFQPEIVISNITENVINLTCSS
HGYPEPKMSVLLRTKNSITIEYDGMOKSDONTVELYDVISLSVSPDVTNNMTIFC
ILETDKTRLLSSPFSIELEDPQPPDHPIHTAVLPTVILCVAVFCILIMKWKKKRP
RNSYKGTNTWREESQTKKREKIHIPERSDEAQRVFKSKTSKSDKSDTCF"

107..175

/gene="B7-2"

176..1093

/gene="B7-2"

/product="CTLA4 counter-receptor"

1112

/gene="B7-2"

sig_peptide

mat_peptide

polyA_site

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1112;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGTGTCTGTGCTAGTCCC 20
|||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 8

AR030780/c 1120 bp DNA linear PAT 29-SEP-1999

LOCUS AR030780

DEFINITION Sequence 1 from patent US 5861310.

ACCESSION AR030780

VERSION AR030780.1 GI:5943994

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1120)

AUTHORS Freeman,G.J., Nadler,L.M. and Gray,G.S.

TITLE Tumor cells modified to express B7-2 with increased immunogenicity

```

and uses therefor
Patent: US 5861310-A 1 19-JAN-1999;
LOCUS
  Location/Qualifiers
  1..1120
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
    |||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 9
ARI12747/c
LOCUS
  Definition
  Sequence 1 from patent US 6130316.
  Accession
  ARI12747
  Version
  ARI12747.1 GI:14092647
  Keywords
  Unknown.
  Source
  ORGANISM
    Homo sapiens (human)
    Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1120)
  Sturmhoeftel, K., Wolf, S.F. and OGTOOLE, M.
  Use of soluble costimulatory molecules to enhance immune responses
  Patent: JP 2002544170-A 2 24-DEC-2002;
  GENETICS INSTITUTE INC
  OS Homo sapiens (human)
  PN JP 2002544170-A/2
  PD 24-DEC-2002
  PF 05-MAY-2000 JP 2000616813
  PR 06-MAY-1999 US 60/132944
  PI KNOT STURMHOFTEL, STANLEY F WOLF, MARGOT O'TOOLE PC
  A61K39/39, A61K39/00, A61P13/02, A61P13/08, A61P15/00, PC
  A61P17/00,
  PC A61P35/00
  CC Use of soluble costimulatory molecules to enhance immune CC
  FH Key
  FT CDS
  Location/Qualifiers
  1..1120
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
    |||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 10
ARI146413/c
LOCUS
  Definition
  Sequence 22 from patent US 6218510.
  Accession
  ARI146413
  Version
  ARI146413.1 GI:15109602
  Keywords
  Unknown.
  Source
  ORGANISM
    Homo sapiens (human)
    Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1120)
  Sharpe, A.H., Borriello, F., Freeman, G.J. and Nadler, L.M.
  B7-1 and B7-2 polypeptides
  Patent: US 6218510-A 22 17-APR-2001;
  Location/Qualifiers
  1..1120
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
    |||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 11
BD272169/c
LOCUS
  Definition
  Sequence 3 from patent US 6352694.
  Accession
  ARI96804
  Version
  ARI96804.1 GI:20246653
  Keywords
  Unknown.
  Source
  ORGANISM
    Homo sapiens (human)
    Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1120)
  June, C.H., Thompson, C.B., Nabel, G.J., Gray, G.S. and Rennert, P.D.
  Methods for inducing a population of T cells to proliferate using
  agents which recognize TCR/CD3 and ligands which stimulate an
  accessory molecule on the surface of the T cells
  Patent: US 6352694-A 3 05-MAR-2002;
  Location/Qualifiers
  1..1120
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
    |||||
Db 57 CGTGTCTGTGCTAGTCCC 38

RESULT 12
ARI96804/c
LOCUS
  Definition
  Sequence 3 from patent US 6352694.
  Accession
  ARI96804
  Version
  ARI96804.1 GI:20246653
  Keywords
  Unknown.
  Source
  ORGANISM
    Homo sapiens (human)
    Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1120)
  June, C.H., Thompson, C.B., Nabel, G.J., Gray, G.S. and Rennert, P.D.
  Methods for inducing a population of T cells to proliferate using
  agents which recognize TCR/CD3 and ligands which stimulate an
  accessory molecule on the surface of the T cells
  Patent: US 6352694-A 3 05-MAR-2002;
  Location/Qualifiers
  1..1120
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTCTGTGCTAGTCCC 20
    |||||
Db 57 CGTGTCTGTGCTAGTCCC 38

```

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 13
LOCUS AR287728 1120 bp mRNA linear PAT 12-JUN-2003
DEFINITION Sequence 3 from patent US 6534055.
ACCESSION AR287728
VERSION AR287728.1 GI:31674734
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1120)
AUTHORS June, C.H., Thompson, C.B., Nabel, G.J., Gray, G.S. and Rennert, P.D.
TITLE Methods for selectively stimulating proliferation of T cells
JOURNAL Patent: US 6534055-A 3 18-MAR-2003;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 14
LOCUS AR374141 1120 bp mRNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6605279.
ACCESSION AR374141
VERSION AR374141.1 GI:40076732
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1120)
AUTHORS Freeman, G.J., Nadler, L.M. and Gray, G.S.
TITLE Therapeutic compositions for inhibiting the interactions of B7-1 and B7-2 with their natural ligands
JOURNAL Patent: US 6605279-A 1 12-AUG-2003;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 15
LOCUS AR381495 1120 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 22 from patent US 6608180.
ACCESSION AR381495
VERSION AR381495.1 GI:40089556
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1120)
AUTHORS Sharpe, A.H., Borriello, F., Freeman, G.J. and Nadler, L.M.
TITLE B7-specific antibodies
JOURNAL Patent: US 6608180-A 22 19-AUG-2003;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 16
LOCUS AR527666 1120 bp mRNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6723705.
ACCESSION AR527666
VERSION AR527666.1 GI:53915455
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1120)
AUTHORS Freeman, G.J., Nadler, L.M. and Gray, G.S.
TITLE Tumor cells modified to express B7-2 with increased immunogenicity and uses therefor
JOURNAL Patent: US 6723705-A 1 20-APR-2004;
FEATURES Location/Qualifiers
source 1..1120
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTGCTGTGCTAGTCCC 20
Db 57 CGTGTGCTGTGCTAGTCCC 38

RESULT 17
LOCUS AX047043 1120 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 3 from Patent WO0067788.
ACCESSION AX047043
VERSION AX047043.1 GI:11876461
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Sturmhoefel, K., Wolf, S.F. and O'Toole, M.
TITLE Use of soluble costimulatory molecules to enhance immune responses
JOURNAL Patent: WO 0067788-A 3 16-NOV-2000;
FEATURES Location/Qualifiers
source 1..1120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
107..1096
/note="unnamed protein product"

CDS

```

/codon_start=1
/protein_id="CAC18947.1"
/db_xref="GI:11876462"
/translation="MDPOCTMGLSNILFVMAFLLSGAAPLKIQAVFNETADLPCCOFAN
SONQSLSELVFWQDENLVNEVYLKKEFDSVHSKYMGRTSPDSKWTLEHLNLOI
KDKGLYQCLIIHKKTGMIRIHMNSELSVLANFQPEIIVPISNITENYINLTGSSI
HGYEPFKMSVLLRTKNSITIEYDGMQSDNVELYDVISLSVSPDVTNMTIFC
ILETDKRLSSPFIIELEDPQPPHPIWITAVLPTVICVMVFLILWKWKKKRRP
RNSYKCGTNTMERBESEQTKREKIHIPERSDEQVFKSKTSKSCDKSDTCF"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CQTGTGTCGTGCTAGTCCC 20
Db 57 CQTGTGTCGTGCTAGTCCC 38
RESULT 18
BC040261/c
LOCUS      BC040261.1      2781 bp      mRNA      linear      PRI 30-JUN-2004
DEFINITION Homo sapiens CD86 antigen (CD28 antigen ligand 2, B7-2 antigen),
transcript variant 1, mRNA (CDNA clone MGC:34413 IMAGE:5173789),
complete cds.
ACCESSION BC040261      GI:25955518
VERSION    BC040261.1
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 2781)
AUTHORS    Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Ziegler,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
Strausberg,R.
2 (bases 1 to 2781)
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/

Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 51 Row: j Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 29029570.

FEATURES
Location/Qualifiers
1..2781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34413 IMAGE:5173789"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2781
/gene="CD86"
/note="Synonyms: B7-2, B70, MGC34413, CD28LG2, LAB72"
/db_xref="LocusID:942"
/db_xref="MIM:601020"
117..1106
/gene="CD86"
/codon_start=1
/product="CD86 antigen, isoform 1"
/protein_id="AAH40261.1"
/db_xref="GI:25955519"
/db_xref="LocusID:942"
/db_xref="MIM:601020"
/translation="MDPOCTMGLSNILFVMAFLLSGAAPLKIQAVFNETADLPCCOFAN
SONQSLSELVFWQDENLVNEVYLKKEFDSVHSKYMGRTSPDSKWTLEHLNLOI
KDKGLYQCLIIHKKTGMIRIHMNSELSVLANFQPEIIVPISNITENYINLTGSSI
HGYEPFKMSVLLRTKNSITIEYDGMQSDNVELYDVISLSVSPDVTNMTIFC
ILETDKRLSSPFIIELEDPQPPHPIWITAVLPTVICVMVFLILWKWKKKRRP
RNSYKCGTNTMERBESEQTKREKIHIPERSDEQVFKSKTSKSCDKSDTCF"
ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 2781;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CQTGTGTCGTGCTAGTCCC 20
Db 67 CQTGTGTCGTGCTAGTCCC 48
RESULT 19
AC068630
LOCUS      AC068630      164161 bp      DNA      linear      PRI 25-NOV-2002
DEFINITION Homo sapiens 3 BAC RP11-289N10 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC068630
VERSION    AC068630.21      GI:16572839
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 164161)
AUTHORS    Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oesman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Briefe,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,K., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

```

Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sleson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 164161)
Worley, K.C.

Direct Submission
Submitted (06-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 164161)
Worley, K.C.

Direct Submission
Submitted (01-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 164161)
Worley, K.C.

Direct Submission
Submitted (25-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 1, 2001 this sequence version replaced gi:16152225.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	source	QUALSTAT-REPORT	Location/Qualifiers
repeat_region	repeat_region	1. .1980	/rpt family="L1MA4"
misc_feature	misc_feature	complement(1. .1972)	/note="overlaps bases 28. .1999 of clone AC068754"
repeat_region	repeat_region	2077. .2651	/function="Clone overlap"
repeat_region	repeat_region	2661. .3430	/rpt family="MER7C"
repeat_region	repeat_region	3433. .3457	/rpt family="L1ME3A"
repeat_region	repeat_region	complement(3463. .3984)	/rpt family="TCTCCC"
repeat_region	repeat_region	complement(3985. .4938)	/rpt family="SVA"
repeat_region	repeat_region	5066. .5492	/rpt family="AluSp"
repeat_region	repeat_region	complement(5575. .5731)	/rpt family="L1MA5"
repeat_region	repeat_region	5732. .6018	/rpt family="AluSp"
repeat_region	repeat_region	6021. .6056	/rpt family="CABA"
repeat_region	repeat_region	complement(6057. .6264)	/rpt family="L1MA5"
repeat_region	repeat_region	7148. .7175	/rpt family="AT-rich"
repeat_region	repeat_region	8832. .9130	/rpt family="L1MA9"
repeat_region	repeat_region	9183. .9383	/rpt family="AluJb"
repeat_region	repeat_region	9384. .9406	/rpt family="CAAAA"
repeat_region	repeat_region	9407. .9421	/rpt family="AluJb"
repeat_region	repeat_region	complement(9789. .10085)	/rpt family="L1ME3"
repeat_region	repeat_region	10150. .10362	/rpt family="MIR"
repeat_region	repeat_region	10895. .10994	/rpt family="CA-rich"
repeat_region	repeat_region	complement(11549. .11625)	/rpt family="L2"
repeat_region	repeat_region	12422. .12704	/rpt family="AluSx"
repeat_region	repeat_region	12705. .12745	/rpt family="CABA"
repeat_region	repeat_region	13035. .13253	/rpt family="MERSA"
repeat_region	repeat_region	complement(13567. .13768)	

```

/rpt_family="MIR"
14327..14640
/rpt_family="ALUSx"
14987..15072
/rpt_family="MIR"
15187..15226
/rpt_family="(TTCA)n"
complement(15227..15332)
/rpt_family="L2"
complement(15339..15539)
/rpt_family="MIR"
complement(16577..16739)
/rpt_family="MIR"
17310..17336
/rpt_family="(GA)n"
17558..17916
/rpt_family="L2"
complement(17921..18090)
/rpt_family="MIR"
18097..18550
/rpt_family="L2"
19164..19451
/rpt_family="L2"
21617..21649
/rpt_family="GC-rich"
21812..21950
/standard_name="91222"
22117..22390
/rpt_family="L2"
22913..23020
/rpt_family="MIR"
complement(23467..23606)
/rpt_family="MIR"

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 178804)
Worley,K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178804)
Worley,K.C.
Direct Submission
Submitted (06-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 6, 2002 this sequence version replaced gi:17646868.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMSC
Center clone name: Rp11-233L3
----- Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 11% of reads
Chemistry: Dye-terminator Big Dye; 89% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 207015 bases at least Q40
Consensus quality: 216378 bases at least Q30
Consensus quality: 223344 bases at least Q20
Estimated insert size: 203159; sum-of-contigs estimation
Quality coverage: 9.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2660: contig of 2660 bp in length
* 2661 2760: gap of unknown length
* 2761 5681: contig of 2921 bp in length
* 5682 5781: gap of unknown length
* 5782 8570: contig of 2789 bp in length

```

* 8571 8670: gap of unknown length
 * 8671 11983: contig of 3313 bp in length
 * 11984 12083: gap of unknown length
 * 12084 17278: contig of 5195 bp in length
 * 17279 17378: gap of unknown length
 * 17379 64030: contig of 46552 bp in length
 * 64031 64130: gap of unknown length
 * 64131 178804: contig of 114674 bp in length.

FEATURES

Location/Qualifiers
 1..178804
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-233L3"

ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 178804;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGTGTCTGTCTGTCTAGTCCC 20
 |||||
 Db 9115 CGTGTCTGTCTGTCTAGTCCC 9134

RESULT 21

AY080604 2710 bp mRNA linear PLN 18-SEP-2002
 Arabidopsis thaliana unknown protein (At3g05380) mRNA, partial cds.

LOCUS

AY080604

ACCESSION

AY080604.1 GI:19347723

VERSION

FLI CDNA.

KEYWORDS

Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 2710)
 Yamada, K., Liu, S.-X., Sakano, H., Pham, P. K., Banh, J., Chung, M. K.,
 Goldsmith, A. D., Lee, J. M., Quach, H. L., Toriumi, M., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayaishizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Nguyen, M., Palm, C. J., Sakurai, T.,
 Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
 Davis, R. W., Ecker, J. R., and Theologis, A.

JOURNAL

Arabidopsis Full Length cDNA Clones

AUTHORS

Unpublished

TITLE

2 (bases 1 to 2710)
 Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
 Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
 Tang, C. C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayaishizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Nguyen, M., Palm, C. J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R. W., Ecker, J. R., and Theologis, A.

JOURNAL

Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA'). Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayaishizaki, Y. and Shinozaki, K.

COMMENT

The Salk, Stanford, PGECC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J.,
 Chan, M. M., Chang, C. H., Chang, E., Dale, J. M., Deng, J. M.,
 Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C.,
 Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
 Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
 Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P.,

Southwick, A., Davis, R. W., Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC)
 contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
 genome submitted to GenBank.

FEATURES

Location/Qualifiers
 1..2710
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL08-15-A12 (R11137)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector
 (FLC-1) as a BamHI/XhoI insert."

gene

<1..2710
 /gene="At3g05380"

CDS

<1..2465
 /gene="At3g05380"
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AAL86288.1"
 /db_xref="GI:19347724"
 /translation="GSPKVESPNRRTSLSDSPISKWGMSTRKSKHCGSSIFE
 EMESSRKLDSKDTLLMDMERAGEMAPRKVKYKRVKVEAECDNDSDNGE
 ACATQGLRSQSRKAAIEASREKYSRPRKRDGHTSGAFDPAQLAELASMLP
 ANMSELSAQLEERTEYDMDEKSTPEATSTSSHGEKANVEPDDLHAISSVENA
 NKRKSPRLVSTDDDVPTGKLQPTSGSLRKRKPKVLGDEAPAEFSQNKSNKKEL
 PDENNMKSLVTKRAGQVPAQSKOMKTVKALRESAITSDKRPGMDIVASPKQVSDS
 GPTLSQKPNRRKSKLQKAEKSETTHKARSRSLSQEALLKDLATSLSP
 PFARRRCIFEFYSADIDHPWFSKMEFVDYLNHVGILHPIRLTRUESVIKSSLRPRR
 FSRFLHEERELKQYRESVRKYHTELRTGAREGLPTDLARPLAVGNRVIAIHPKTR
 IHGKILITVDHNKCNVLFDDLGVELVMDIDCMPLNPLEYMEGLRQIDKLSMKEA
 QLSGNTNLGVSLFPFCGLNVFSNMNPLNQGDMAPIILHGVSSNTSPQTNHSY
 IITYNKAEAEIORAQALHDLDEKEMPEMLIVKSKTRQAQAMVDAAIKAASVKE
 GEDVNTMQALELVGNKQLLRSMVYKHVNGSIEHHNPSNGSEPVANNDLNS
 QDGEKNAQMPSELITSCVATVLMIQMCTERQYPPADVQLIDIANVTSQPRCPQNLFP
 IYRIEQTCMGRIKTIQINSLVPT"
 2466..2710
 /gene="At3g05380"
 /misc_difference 2694
 /gene="At3g05380"
 /note="not present in genomic sequence"

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 2710;
 Best Local Similarity 94.7%; Pred. No. 4.8e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GTGTCTGTCTGTCTAGTCCC 20
 |||||
 Db 1726 GTGTCTGTCTGTCTATCCC 1744

RESULT 22

CQ803940 3168 bp DNA linear PAT 10-MAY-2004
 LOCUS CQ803940

DEFINITION

Sequence 351 from Patent WO2004035798.

ACCESSION

CQ803940

VERSION

CQ803940.1 GI:47110604

KEYWORDS

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

AUTHORS

Inze, D., de Veylder, L. and Vlieghe, K.

TITLE

Identification of novel e3f target genes and use thereof

JOURNAL Patent: WO 2004035798-A 351 29-APR-2004;
CropDesign N.V. (BE)
FEATURES Location/Qualifiers
source 1. 3168
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 87.0%; Score 17.4; DB 6; Length 3168;
Best Local Similarity 94.7%; Pred. No. 4.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20
|||||
Db 2429 GTGTCTCTGCTATGCC 2447

RESULT 23
AJ583496
LOCUS 3518 bp mRNA linear PLN 22-JUL-2004
DEFINITION Arabidopsis thaliana mRNA for Arabidopsis always early 2 protein (Ataly2 gene).
ACCESSION AJ583496
VERSION AJ583496.1 GI:50539419
KEYWORDS always early 2 protein; Ataly2 gene.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Bhatt,A.M., Zhang,Q., Harris,S.A., White-Cooper,H. and Dickinson,H.
TITLE Gene structure and molecular analysis of Arabidopsis thaliana ALWAYS EARLY homologs
PUBMED 15246533
REFERENCE 2 (bases 1 to 3518)
AUTHORS Bhatt,A.M.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2003) Bhatt A.M., Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, UNITED KINGDOM

FEATURES source Location/Qualifiers
1. 3518
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/tissue_type="inflorescence"
/note="ecotype: Columbia"
1. 3518
/gene="aly2"
135. 3290
/gene="aly2"
/codon_start=1
/evidence=experimental
/product="always early 2 protein"
/protein_id="CAE47461.1"
/db_xref="GI:50539420"
/translation="MAPVKRSVNNKRFNTSPKADGKSKNKLKSLDKLGPWG TRLEUFRFDYAKHGQSWRRVAARINRSVDVMEALFNMNRAVLSLPEGTASVAGL LAMTMDHYSVMEGSGEGHDASVPKQKQKAPKQSDSPREVDIQOISGPDG CLTFIKQARANGTORHATGKTPRPVQVTSFNRDDREGSTPNKARKQFPANDVAH FIALATDASRRGSPKVSSEPNRTELSDDSPISKWGMKSRKSKQKHCSSIFEE WMESRRKLSDKDTLLMDMERAGMEAPKGRVYKRVKVEEACNDSDDNGE CSATQGLRSQRKAAAEAREKYSRPSKPKRDDKHTSGAFDALQALAEIASMLPA NMESLGAQLKEERTYDMDKSTPATSTSHCEKANVPDPSLLHATISSVBNAN KRKSPRLVSTCDVPTGKLPQTSGLRKRKPKVLGDEAPAFSONKSKINKELP QDENNMKSLVTKRAGQPAQSKQMTVKALEESAITSDKKRPGCHDI VASPKQVSDSG PTLUSQKPPNRRKSLQKLSQKAEKSSSTTHKAARSSLSGELELLKDLKATLSLSPF FARRRCIFEWYSADHPWFSKMEFVYLNHVLGHIPLRLTLEWSVSKSLGRPRF SERFLHEERELKQYRESVRKHYTELRTGAREGLPTDLARPLAVGNRVIAHPKTREI

FEATURES
Location/Qualifiers

HDGKILTVDHNNKNVFLDGLGVLMIDCMPLNPLEYMPBGLRRQIKDKLSMKKEAQ
LSGNTNLGVSILFPFCGLENSYFSNPNPILNQDMIAPIILHGKVSNTSPSPQTNHSYI
TTYNKAKAEIQRQAALQHALDEKEMEPEMLEIKVGSQTRAAQAMVDAAIKAASSVKEG
EDVNTMIOEALVELKGNLSSVMVKKHHEVNGSIEHHNPSNGSPVANNDLNSQ
DGSEKQMPSELATSCVATWLMQMCITERQYPPADVAQLIDAAVTSIQPRCPQLPI
YREIQTCMGRIKTIQIMSLVPI"

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 3518;
Best Local Similarity 94.7%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGCTAGTCCC 20
|||||
Db 2551 GTGTCTCTGCTATGCC 2569

RESULT 24
ATAC009606/c
LOCUS 91924 bp DNA linear PLN 30-OCT-2002
DEFINITION Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence, complete sequence.
ACCESSION AC009606
VERSION AC009606.4 GI:12484386
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 91924)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B., Wu,D., Maiti,R., Rensing,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
TITLE Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence Unpublished
JOURNAL 2 (bases 1 to 91924)
AUTHORS Lin,X. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
REFERENCE 3 (bases 1 to 91924)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT On Jan 25, 2001 this sequence version replaced gi:12280792.
Address all correspondence to:atetigr.org

BAC clone F22F7 is from Arabidopsis chromosome III and is near the molecular marker mil72.
The orientation of the sequence is from SP6 to T7 end of the BAC clone.
Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene (http://www.cbs.dtu.dk/netpGene/cbsnetpGene.html), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (http://www.tigr.org/tdb/at.html).
Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by RepeatMasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).

FEATURES
Location/Qualifiers

Query Match

Best Local Similarity 87.0%; Score 17.4; DB 8; Length 91924;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTCTAGTCCC 20

DB 76048 GTGTCTCTGTCTATCCC 76030

RESULT 25

AC140972/c 144858 bp DNA linear HTG 29-MAY-2003

LOCUS Felis catus clone RP86-144L3, WORKING DRAFT SEQUENCE, 11 ordered
 DEFINITION pieces.

ACCESSION AC140972

VERSION AC140972.2 GI:31126657

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE Felis catus (cat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE 1 (bases 1 to 144858)

AUTHORS

Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
 Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
 Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
 Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O.,
 Legaspi,R., McDowell,J., Maduro,V.B., Margulies,E.H., Mastello,C.,
 Maskeri,B., Pearson,C., Pearson,R.K., Portnoy,W.E.,
 Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
 Sison,C., Standtropol,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
 Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

TITLE

NISC Comparative Sequencing Initiative

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 144858)
 Green,E.D.
 Direct Submission
 Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

JOURNAL

3 (bases 1 to 144858)

REFERENCE

Green,E.D.
 Direct Submission
 Submitted (29-MAY-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

JOURNAL

On May 29, 2003 this sequence version replaced gi:28867004.

COMMENT

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@hgri.nih.gov
 ----- Project Information
 Center project name: efa
 Center clone name: 144L03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 142769 bases at least Q40
 Consensus quality: 143488 bases at least Q30

Consensus quality: 143725 bases at least Q20
 Insert size: 168000; agarose-fp
 Insert size: 143858; sum-of-contigs
 Quality coverage: 10.45x in Q20 bases; agarose-fp
 Quality coverage: 12.20x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 908: contig of 908 bp in length
 * 909 1008: gap of unknown length
 * 1009 5899: contig of 4891 bp in length
 * 5900 5999: gap of unknown length
 * 6000 17014: contig of 11015 bp in length
 * 17015 17114: gap of unknown length
 * 17115 19195: contig of 2081 bp in length
 * 19196 19295: gap of unknown length
 * 19296 78277: contig of 58982 bp in length
 * 78278 78377: gap of unknown length
 * 78378 85051: contig of 6674 bp in length
 * 85052 85151: gap of unknown length
 * 85152 87353: contig of 2202 bp in length
 * 87354 87453: gap of unknown length
 * 87454 89662: contig of 2109 bp in length
 * 89663 125083: contig of 35421 bp in length
 * 125084 125183: gap of unknown length
 * 125184 13715: contig of 12432 bp in length
 * 13716 137715: gap of unknown length
 * 137716 144858: contig of 7143 bp in length.

FEATURES

source

1..144858
 /organism="Felis catus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9685"
 /clone="RP86-144L3"
 /clone_lib="RP86"

misc_feature

1..908
 /note="assembly_fragment"
 clone_end:77
 vector_side:left"

misc_feature

1009..5899
 /note="assembly_fragment"

misc_feature

6000..17014
 /note="assembly_fragment"

misc_feature

17115..19195
 /note="assembly_fragment"

misc_feature

19296..78277
 /note="assembly_fragment"

misc_feature

78378..85051
 /note="assembly_fragment"

misc_feature

85152..87353
 /note="assembly_fragment"

misc_feature

87454..89562
 /note="assembly_fragment"

misc_feature

89663..125083
 /note="assembly_fragment"

misc_feature

125184..137615
 /note="assembly_fragment"

misc_feature

137716..144858
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 144858;
 Best Local Similarity 94.7%; Pred. No. 2.2e+02;

Matches	18; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	1	CGTGTGCTGCTAGTCC	19		
Db	131485	CGTGTGCTTGTAGTCC	131467		
RESULT 26					
LOCUS	AC136223	146492 bp	DNA linear	PLN 09-AUG-2004	
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 5 clone				
ACCESSION	OSJNBa0093E24, complete sequence.				
VERSION	AC136223				
KEYWORDS	HTG.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.				
AUTHORS	1 (bases 1 to 146492)				
	Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R., Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Liu, H.-L., Li, Y.-P., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y., Yu, S.-W., Wu, H.-P., Shaw, J.-F., *McCombie, W.R., *de la Bastide, M., *Spiegel, L., *Zutavern, T., *Muller, S., *Nascimento, L., *Baliia, V., *Bell, M., *Miller, B., *Katzenberger, F., *Andrade, M.V., *Dike, S., *O'Shaughnessy, A., *Palmer, L. and *Dedhia, N.				
TITLE	Oryza sativa (japonica cultivar-group) chromosome 5 clone				
JOURNAL	OSJNBa0093E24, complete sequence				
REMARK	*Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA				
AUTHORS	2 (bases 1 to 146492)				
TITLE	Chow, T.-Y. and Hsing, Y.-I. C.				
JOURNAL	Direct Submission				
AUTHORS	Submitted (31-OCT-2002) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan				
TITLE	3 (bases 1 to 146492)				
JOURNAL	Chow, T.-Y. and Hsing, Y.-I. C.				
AUTHORS	Direct Submission				
TITLE	Submitted (25-FEB-2004) Institute of Botany, Academia Sinica, 128, Section2, Academia Road, Nankang, Taipei 11529, Taiwan				
JOURNAL	4 (bases 1 to 146492)				
AUTHORS	Chow, T.-Y. and Hsing, Y.-I. C.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-AUG-2004) Institute of Botany, Academia Sinica, 128, Section2, Academia Road, Nankang, Taipei 11529, Taiwan				
COMMENT	On Feb 25, 2004 this sequence version replaced gi:24431494. This BAC clone was sequenced to phase II by the Academia Sinica and was completed to phase III by the Cold Spring Harbor Laboratory Genome Center. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. Genes were predicted from the integrated results of the following: BLASTN2.0, BLASTX2.0, GENSCAN (Chris Burge, http://genes.mit.edu/GENSCAN.html), Fgenesh (http://www.softberry.com/), GlimmerK (http://www.tigr.org/softlab/glimmer/glimmer.html), TWINSKAN (http://genes.cs.wustl.edu/) and GeneSplicer (http://www.tigr.org/tdb/GeneSplicer/index.shtml). The sequence was searched against the Swiss-Prot-TrEMBL protein database, the NCBI Plant EST database, the TIGR Rice Gene Index and the rice full-length cDNA database (KOME, http://cdna01.dna.affrc.go.jp/cDNA/). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits.				

```
/db_xref="UniProt/Trembl:Q84SJS"
/translation="MTKPSKTEVEPSNMPVTLDDFEGEDRKAMEEYINELTRBALM
RATTRTQCVIIPKPRPKLADPKVDKDFGLNRLGSLMDDYTLKDKAYTANPPID
QTSKTDGTAFTAGTDPGRDRTAVGTQSGTQGTGYPGQGTGYSQAGTAHDA
GOTSGRAGGTATVIOAIADPMTNASLYVHFTFDPVSTATSNQVPHIIPNAKEA
SRGYPDTRHGQNHIAPTQPIRPPLPNPQGNMEDMISDVIRKFGLETNRRA
RTYKTIPIYDYNPPQYRVPEFQKSGEDGKTTWEHQFLAQCEAGNSDTPKJR
LFSLSNGAFTWFTSLPANSIHTWAQEEQPHDYFAGETELRCLDTSVKQKYES
VIDYIKRFRDVRNQCVTLMITDRDLAGLAFNGLVALLRDLRGLDQFLDVSQMLQAL
QBSRVKDKLAPDPKKNVNVVDYPKASDSSEEDGHDIIYVAEWSWTKNPKPVCNS
LMPTRKQOSKRVFQVAKCDKIDYLLQEKQJRLPKGHVIPSQELKRAYCKWH
DSHSTNDVNFVRQVQSAIDEGRLKFTDGSKMLDHDPPVPTNFDNKKVLIRPE
QASTKGVNIGEPSEMIPEKPEVGVKKNRSGASTSKAPRTKTKTDFMLLEKIE
KQGEKRGKRVERTSEVITIKVGHSDVPTIPSGDVGKSPKSEAGTSSSAGIL
TRPFGSDCGHAAGTDPGSDRDPDGTDFGRFDRFVRPHVAGLIGLQRRFDRIV
KGDVRASSLPVMPKSTLEKAPIDIVEEKESQSDKSESSPKEDMDINMFCAMDEAE
RDKRNELOPPVMPKSTLEKAPIDIVEEKESQSDKSESSPKEDMDINMFCAMDEAE
VAFSLGPKDAVPEKDESNRMKFLYLKGVHDGKPVSRMLVDGGAANVLMPSYLFKK
KREDEDELUKTNILNGEPTFTEARGIFASBLTVGNKTLPTAFPIVDVQEANLEWLQ
KRVQYRSTKNDIGETDFDEKLGQGLTADPLEVDIGDFTKPRPTFVNKNMRA
DYKVKIIEILLKAGADFCEWYHEIPLGSELVHEHFPPIKSGFRPKQPPHFNPLLYD
RVKEIDILLKAGDFPCRYAEWSSIVPEKESKIRVCIDFRLNKAATPKDREYPM
PIADMDINDASGHKVISLGDAGYNOIPEAEEDMYKTAFCPCFVGLFEWVMTFGL
KNAGATYQALNIEFHLLGLVLEIYIDIVVKSDEADTFGAKQOEAFEEELKRYLST
PPVRAKAGKFRILYIASEDKVIGAVLTQEDGKEYITILSRLLDAETRPILSGR
IGKWAYALIEYDAYEPLKSMKGQIVCFPIIDHIDIDCAVEGEVLEIVEIPWKIYFNGS
SCKEGGIVGLVFPNGMCEYASARLEYCTNNOAYNALLPGLQVMEVMAKYVEAF
GDSLVQVQVAGVYKCLDLSNLYLSDCLDIITANFDNPAIRHIAARDNSRSDWPHATGL
SSVNVKRGFLJLEJPELSDRLCEIGKIRDOGRSDRHTAGLTGNQSGSDWPHATGL
TGNPERSNRCVVGITGSGQAGHVSINLEABLIVNSDVCAQETEGEDWRIPLIQTLLK
DPTLKVDRIKQAFKTYLLDGDLYRNIIDGVLKCLDGLSKSANGSGHREVFVATD
YFTKWAELVIRPQYTHTEANGAESNTLLKLVKIKIYEPKRVHEVLSALMAHRI
SKHGAIRVFPFELAYQGVAVLPEVNLGSLHIIKQDDLSSEYKMLMGNDLDEVIDDKR
LKALREIEKKEKRVKAKYKAKLVQVGLDVAWTKTILPLGTRSEKFGKWSFSGEYPI
RVCGIVRGVNAVYFLEALQGERFORAINGKVLKYPFVWQDA"
complement(38822..43272)
/gene="OSJNBa009324.5"
/feature="putative polyprotein"
complement(join(38822..39873,39931..40071,40154..40955,
41202..41582,41663..41756,41762..43272))
/gene="OSJNBa009324.5"
/codon_start=1
/product="putative polyprotein"
/protein_id="AAT93986.1"
/db_xref="GI:51038183"
/db_xref="UniProt/Trembl:Q7XFK3"
/translation="MNNANRKGFGNRCFEGCSIDLHLSHCPKLGKRGKREKQGEKT
NNKNNNNKSGLSQGRKLDNRKAPQQVCAVFEPLSDVDSGGDDDKGKNSIDVCFM
ARGEDTYEDNEVSAFEAEINILSAKKKYEYRQKEFTIESLKSIDRLKSLIPN
DGCSCVEVLMNEITKIRDVNAHDLNRRSLACNFALHTRTLDLFLAKLLQKYQI
AFHSLMFEMI SAKLLKOPHDVLDCTCNLNMKLDALGRVYEMDVYKNNVLSQP
KCKSKGVMDCCENCANLEKVSYLNNILRESEKGNLMLLDQSKVSTNHRGFPN
PYAHSRHPVVLGVGARSCEILLVPTDNTKTVFKSGGIMSTHAKSSSSNVVHAKSPV
VASAKSSVSNVNNREKTCSCFGKHGIVGFCFLAHKOKKREIEAFKSKWQKS
GYSRKPFRPQWLVSDRVPVATERRSDQAPVAVQKSNRPVPPVATVPREPATV
PRSDRPRNSFENLILHVIDICLLDLVALVARKENNIWVDSGCSHITGDKNWFSSLK
KASKTENPEFPKTRKSVDFSDGLVNLISRYGRVFKADPENPVSPVITCLVAKFDK
DVMFWHRLGHVGHDLTLELSDLIIRDLPLKDDLDLEPPVSLKIRSDNGGEFK
NASFQCNERGLEHFSRVPVQNGVVRKRVILVEMATMLVEYTKPRKDFPKKI
HHFPHKTSYELRFGHQPKVSLRVFGCKFVLKSRNLDKFEARSTNGLFLGYPATHH
GYRVLGTNKIIVETCEVDSFSGTRPDIACTISQVQGEDGHTFDEGDNNDDE
VGSAGTQCAQTAGTSVRPAQVRSDRPGSAEGTVADRDGPETITTSVAAPL
HIQRHHPPEQIIGNIDVSHALTDESWINAMKELENFERNTVMTLVEPPPSGHNIGTK
```

Query Match 87.0%; Score 17.4; DB 8; Length 146492;
Best Local Similarity 94.7%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 GTGTGCTGTGCTAGTCCC 20
DB 104816 GTGTGCTGTGCTACTCCC 104834

RESULT 27
CR391906
LOCUS
DEFINITION
Danio rerio clone DKEY-211K10, *** SEQUENCING IN PROGRESS ***, 9
unordered pieces.
ACCESSION
CR391906.2 GI:46559615
VERSION
HTG; HTGS; PHASE1.
KEYWORDS
Danio rerio (zebrafish)
SOURCE
Danio rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
Mclay, K.
Direct Submission
Submitted (23-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 24, 2004 this sequence version replaced gi:46517964.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zk211K10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 159202 bases at least Q40
Consensus quality: 159648 bases at least Q30
Consensus quality: 160053 bases at least Q20
Insert size: 160716; sum-of-contigs
Insert size: 167493; 3.3% error; agarose-fp
Quality coverage: 6.27x in Q20 bases; sum-of-contigs Quality
coverage: 6.01x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2483: contig of 2483 bp in length
* 2484 2583: gap of 100 bp
* 2584 38899: contig of 36316 bp in length
* 38900 38999: gap of 100 bp
* 39000 55041: contig of 16042 bp in length
* 55042 55141: gap of 100 bp
* 55142 58203: contig of 3062 bp in length
* 58204 58303: gap of 100 bp
* 58304 70412: contig of 12109 bp in length
* 70413 70512: gap of 100 bp
* 70513 100144: contig of 29632 bp in length
* 100145 100244: gap of 100 bp
* 100245 110317: contig of 10073 bp in length
* 110318 110417: gap of 100 bp
* 110418 122844: contig of 12427 bp in length
* 122845 122944: gap of 100 bp
* 122945 161516: contig of 38572 bp in length.
Location/Qualifiers
1..161516
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-211K10"
/clone_lib="DanioKey"
1..2483
/note="assembly fragment:00070
fragment_chain:1"

misc_feature

```

misc_feature 2584..38899
              /note="assembly fragment:01225
              fragment chain:1"
misc_feature 39000..55041
              /note="assembly fragment:00464
              fragment chain:1"
misc_feature 55142..58203
              /note="assembly fragment:00087
              fragment chain:1"
misc_feature 58304..70412
              /note="assembly fragment:00114
              fragment chain:2"
misc_feature 70513..100144
              /note="assembly fragment:00640
              fragment chain:2"
misc_feature 100245..110317
              /note="assembly fragment:00215
              fragment chain:3"
misc_feature 110418..122844
              /note="assembly fragment:00326
              fragment chain:3"
misc_feature 122945..161516
              /note="assembly fragment:00891.0"

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 161516;
Best Local Similarity 94.7%; Pred No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTCTCTGTGCTAGTCCC 20
   |||||
Db 50287 GTGTCTCTGTGCTAGTCCC 50305

RESULT 28
AC129160
LOCUS
DEFINITION Rattus norvegicus clone CH230-108G14, *** SEQUENCING IN PROGRESS
ACCESSION AC129160
VERSION AC129160.3 GI:30580645
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 253259)
          Muzny,D,Marie,, Metzker,M,Lee,, Abramzon,S,, Adams,C,, Alder,J,,
          Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,,
          Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
          Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
          Biswalo,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
          Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,,
          Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,,
          Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,
          Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Souza,L,,
          Davila,M.L., Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,,
          Delgado,O., Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,,
          Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
          Egan,A,, Escotto,M,, Eugene,C,, Evans,C.A., Falls,T., Fan,G,,
          Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
          Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
          Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K,,
          Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,
          Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
          Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
          Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
          Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
          Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
          Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
          Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,

```

```

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajic,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,Z., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
2 (bases 1 to 253259)
Worley,K.C.
Direct Submission
Submitted (27-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253259)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23111082.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLEN
Center clone name: CH230-108G14
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 234406 bases at least Q40
Consensus quality: 237952 bases at least Q30
Consensus quality: 240010 bases at least Q20
Estimated insert size: 246479; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 4024: contig of 4024 bp in length
 * 4025 4124: gap of unknown length
 * 4125 248130: contig of 244006 bp in length
 * 248131 248230: gap of unknown length
 * 248231 251975: contig of 3745 bp in length
 * 251976 252075: gap of unknown length
 * 252076 253259: contig of 1184 bp in length.

FEATURES

source
 1. .253259
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-108G14"
 misc_feature
 1. .1774
 /note="wgs_contig"
 misc_feature
 4125. .5168
 /note="wgs_contig"
 misc_feature
 5219. .7303
 /note="wgs_contig"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 253259;
 Best Local Similarity 94.7%; Pred. No. 2e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGTCGTGCTAGTCCC 20

Db 130398 GTGTGTCGTGCTAGTCCC 130416
 |||||

RESULT 29

AC094647/c

LOCUS Rattus norvegicus clone CH230-4N1, WORKING DRAFT SEQUENCE, 7
 DEFINITION linear HTG 09-MAY-2003

AC094647 6 GI:30466422

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 261391)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, J.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, M., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhera, L., Loulseged, H., Lozado, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nye, L.,
 Nwaokemele, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Shen, H.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Tatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 261391)

Worley, K. C.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL</

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 250290: contig of 250290 bp in length
 * 250291 250390: gap of unknown length
 * 250391 252114: contig of 1724 bp in length
 * 252115 252214: gap of unknown length
 * 252215 253404: contig of 1190 bp in length
 * 253405 253504: gap of unknown length
 * 253505 254576: contig of 1072 bp in length
 * 254577 254676: gap of unknown length
 * 254677 256299: contig of 1623 bp in length
 * 256300 256399: gap of unknown length
 * 256400 257642: contig of 1243 bp in length
 * 257643 257742: gap of unknown length
 * 257743 261391: contig of 3649 bp in length.

FEATURES

source
 1. 261391
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-4N1"

misc_feature

1. 1302
 /note="wgs end extension
 clone_end:5p6"
 complement(7548..8366)
 /note="clone boundary
 clone_end:5p6
 site:EcoRI
 end sequence:BH309507"

misc_feature

248018..248389
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH309505"

misc_feature

248563..250290
 /note="wgs end extension
 clone_end:T7"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 261391;
 Best Local Similarity 94.7%; Pred. No. 2e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTGTCTGTGTAGTCC 20

Db 161920 GTGTCTGTGTAGTCC 161902

RESULT 30

AC129648/c

LOCUS Rattus norvegicus clone CH230-8H18, WORKING DRAFT SEQUENCE, 3
 DEFINITION linear HTG 13-NOV-2002

AC129648 unordered pieces.

AC129648.4 GI:24940664

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 280215)

REFERENCE

AUTHORS

Muzny,D,Marie, Metzker,M, Lee, Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Behnam,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kwis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenshewala,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwakoeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L.,
 Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Richards,S., Riggs,F.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snajls,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,D., Tingle,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 280215)

Worley,K.C.

Direct Submission

Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 280215)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:23096222.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

```

----- Project Information
Center project name: GDWJ
Center clone name: CH230-8H18
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 237591 bases at least Q40
Consensus quality: 239484 bases at least Q30
Consensus quality: 240656 bases at least Q20
Estimated insert size: 244055; sum-of-coverage estimation
Quality coverage: 7x in Q20 bases; sum-of-coverage estimation
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 255140: contig of 255140 bp in length
* 255141 255240: gap of unknown length
* 255241 274427: contig of 19187 bp in length
* 274428 274527: gap of unknown length
* 274528 280215: contig of 5688 bp in length.
*
* Location/Qualifiers
* 1..280215
*   /organism="Rattus norvegicus"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:10116"
*   /clone="CH230-8H18"
*   1..1128
*   /note="wgs contig"
*   1179..5584
*   /note="wgs_contig"
*
* misc_feature
* misc_feature
*
* ORIGIN
*
Query Match      87.0%; Score 17.4; DB 2; Length 280215;
Best Local Similarity 94.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGTGCTCTGTCTAGTCC 20
    |||||
Db 135089 GTGTGCTCTGTCTAGTCC 135071

RESULT 31
AC113196/c
LOCUS          24804 bp      DNA      linear      HTG 26-FEB-2002
DEFINITION    Mus musculus clone RP23-172K11, LOW-PASS SEQUENCE SAMPLING.
ACCESSION     AC113196
VERSION       AC113196.1 GI:18921382
KEYWORDS      HTG; HGSC PHASE0.
SOURCE        Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              1 (bases 1 to 24804)
REFERENCE     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
              Unpublished
              2 (bases 1 to 24804)
REFERENCE     Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
              Brown,A., Camarata,J., Campopiano,A., Chang,J.J., Chazaro,B.,
              Choepel,Y., Collangelo,M., Collins,S., Collumore,A., Cook,A.,
              Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
              Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
              Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
              Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
              Landers,T., Lehoczký,J., Levine,R., Liu,G., MacLean,C.,

```

```

Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L23475
Center clone name: 172_K_11
-----
* NOTE: This record contains 31 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 716: contig of 716 bp in length
* 717 816: gap of 100 bp
* 817 1520: contig of 704 bp in length
* 1521 1620: gap of 100 bp
* 1621 2327: contig of 707 bp in length
* 2328 2427: gap of 100 bp
* 2428 3149: contig of 722 bp in length
* 3150 3249: gap of 100 bp
* 3250 3968: contig of 719 bp in length
* 3969 4068: gap of 100 bp
* 4069 4762: contig of 694 bp in length
* 4763 4862: gap of 100 bp
* 4863 5567: contig of 705 bp in length
* 5568 5667: gap of 100 bp
* 5668 6345: contig of 678 bp in length
* 6346 6446: gap of 100 bp
* 6446 7147: contig of 702 bp in length
* 7148 7247: gap of 100 bp
* 7248 7966: contig of 719 bp in length
* 7967 8066: gap of 100 bp
* 8067 8741: contig of 675 bp in length
* 8742 8841: gap of 100 bp
* 8842 9561: contig of 720 bp in length
* 9562 9661: gap of 100 bp
* 9662 10355: contig of 694 bp in length
* 10356 10455: gap of 100 bp
* 10456 11377: contig of 682 bp in length
* 11378 11237: gap of 100 bp
* 11238 11931: contig of 694 bp in length
* 11932 12031: gap of 100 bp
* 12032 12747: contig of 716 bp in length
* 12748 12847: gap of 100 bp
* 12848 13532: contig of 685 bp in length
* 13533 13632: gap of 100 bp
* 13633 14312: contig of 680 bp in length

```

TITLE
JOURNAL

COMMENT

* 14313 14412: gap of 100 bp
* 14413 15104: contig of 692 bp in length
* 15105 15204: gap of 100 bp
* 15205 15922: contig of 718 bp in length
* 15923 16022: gap of 100 bp
* 16023 16744: contig of 722 bp in length
* 16745 16844: gap of 100 bp
* 16845 17550: contig of 706 bp in length
* 17551 17650: gap of 100 bp
* 17651 18354: contig of 704 bp in length
* 18355 18454: gap of 100 bp
* 18455 19178: contig of 724 bp in length
* 19179 19278: gap of 100 bp
* 19279 19589: contig of 711 bp in length
* 19590 20089: gap of 100 bp
* 20090 20807: contig of 718 bp in length
* 20808 20907: gap of 100 bp
* 20908 21601: contig of 694 bp in length
* 21602 21701: gap of 100 bp
* 21702 22396: contig of 695 bp in length
* 22397 22496: gap of 100 bp
* 22497 23210: contig of 714 bp in length
* 23211 23310: gap of 100 bp
* 23311 24003: contig of 693 bp in length
* 24004 24103: gap of 100 bp
* 24104 24804: contig of 701 bp in length.

FEATURES

source
1..24804
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-172K11"
/clone_lib="RPCL-23 Female Mouse BAC"

ORIGIN

Query Match 85.0%; Score 17; DB 2; Length 24804;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GTGTCTGTGCTAGTCCC 20
|||||

Db 4721 GTGTCTGTGCTAGTCCC 4705

RESULT 32

LMFLCHR15_6

WPCOMMENT

Sequence split into 8 fragments LOCUS LMFLCHR15 Accession AL160371

Fragment Name	Begin	End
LMFLCHR15_0	1	110000
LMFLCHR15_1	100001	210000
LMFLCHR15_2	200001	310000
LMFLCHR15_3	300001	410000
LMFLCHR15_4	400001	510000
LMFLCHR15_5	500001	610000
LMFLCHR15_6	600001	710000
LMFLCHR15_7	700001	720704

Continuation (7 of 8) of LMFLCHR15 from base 600001 (AL160371 Leishmania major chromosome)

Query Match 85.0%; Score 17; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TGTGTCTGTGCTAGTCCC 19
|||||

Db 52209 TGTGTCTGTGCTAGTCCC 52225

RESULT 33

CR382128_04/c

WPCOMMENT

Sequence split into 31 fragments LOCUS CR382128 Accession CR382128

Fragment Name	Begin	End
---------------	-------	-----

CR382128_00	1	110000
CR382128_01	100001	210000
CR382128_02	200001	310000
CR382128_03	300001	410000
CR382128_04	400001	510000
CR382128_05	500001	610000
CR382128_06	600001	710000
CR382128_07	700001	810000
CR382128_08	800001	910000
CR382128_09	900001	1010000
CR382128_10	1000001	1110000
CR382128_11	1100001	1210000
CR382128_12	1200001	1310000
CR382128_13	1300001	1410000
CR382128_14	1400001	1510000
CR382128_15	1500001	1610000
CR382128_16	1600001	1710000
CR382128_17	1700001	1810000
CR382128_18	1800001	1910000
CR382128_19	1900001	2010000
CR382128_20	2000001	2110000
CR382128_21	2100001	2210000
CR382128_22	2200001	2310000
CR382128_23	2300001	2410000
CR382128_24	2400001	2510000
CR382128_25	2500001	2610000
CR382128_26	2600001	2710000
CR382128_27	2700001	2810000
CR382128_28	2800001	2910000
CR382128_29	2900001	3010000
CR382128_30	3000001	3066374

Continuation (5 of 31) of CR382128 from base 400001 (CR382128 Yarrowia lipolytica chromosome)

Query Match 85.0%; Score 17; DB 8; Length 110000;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GTGTCTGTGCTAGTC 18
|||||

Db 91229 GTGTCTGTGCTAGTC 91213

RESULT 34

BC076559/c

LOCUS

DEFINITION

BC076559 2575 bp mRNA linear VRT 27-JUL-2004

Danio rerio zgc:92474, mRNA (cdna clone MGC:92474 IMAGE:7044589),

complete cds.

ACCESSION BC076559.1 GI:49901363

VERSION BC076559

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 2575)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A.C., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2575)
 Strausberg, R.
 Direct Submission
 Submitted (06-JUL-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@hgrl.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.B., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 178 Row: p Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.
 Location/Qualifiers
 1..2575
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:92474 IMAGE:7044589"
 /tissue_type="Whole, adult zebrafish"
 /clone_lib="NIH ZGC_10"
 /lab_host="DH10B"
 /note="Vector: pExpress1"
 1..2575
 /gene="zgc:92474"
 /db_xref="LocusID:436917"
 156..575
 /gene="zgc:92474"
 /codon_start=1
 /product="zgc:92474"
 /protein_id="AAH76559.1"
 /db_xref="GI:49901364"
 /db_xref="LocusID:436917"
 /translation="MAQVKSIAQSLYPCSQRLHOMKLIQKNSVTCNDGTGAGYY
 IKESGRRLVFLGGWYCFKSHYCDRSYEMRLMSSNPPTRTGTGILSPQPEE
 NPHWMNANTVFYPCSSDVSWSGSTPKTDQRHQGALN"
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 5; Length 2575;
 Best Local Similarity 90.0%; Pred. No. 9.8e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CCGTGTCTGTGCTAGTCC 20
 |||||
 Db 101 CCGTGTGTGTGCTGTGCTCC 82
 |||||
 RESULT 35
 Query Match 84.0%; Score 16.8; DB 5; Length 2575;
 Best Local Similarity 90.0%; Pred. No. 9.8e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CCGTGTCTGTGCTAGTCC 20
 |||||
 Db 101 CCGTGTGTGTGCTGTGCTCC 82
 |||||

AF220172S2/c
 LOCUS Homo sapiens acid ceramidase (ASAH) gene, exons 2 through 4.
 DEFINITION Homo sapiens acid ceramidase (ASAH) gene, exons 2 through 4.
 ACCESSION AF220173 AF220174
 VERSION AF220173.1 GI:9651700
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 14235)
 AUTHORS Mukherjee, A.B.
 TITLE Human acid ceramidase gene
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 14235)
 AUTHORS Mukherjee, A.B.
 TITLE Direct Submission
 JOURNAL Submitted (29-DEC-1999) HDB, NICHD/NIH, 9000 Rockville Pike,
 Bethesda, MD 20892-1830, USA
 FEATURES
 Location/Qualifiers
 1..14235
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8p22-21.2"
 6302..6348
 /gene="ASAH"
 /number=2
 10518..10608
 /gene="ASAH"
 /number=3
 12035..12121
 /gene="ASAH"
 /number=4
 exon
 exon
 exon
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 14235;
 Best Local Similarity 90.0%; Pred. No. 7e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CCGTGTCTGTGCTAGTCC 20
 |||||
 Db 5757 CCGTGTCTGTGCTATTCCC 5738
 |||||
 RESULT 36
 AC118055
 LOCUS Homo sapiens chromosome 15 clone CTD-2013P3 map 15, LOW-PASS
 DEFINITION SEQUENCE SAMPLING.
 AC118055
 VERSION AC118055.1 GI:20143784
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 58969)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 15, clone CTD-2013P3
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 58969)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarato, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collumore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, D.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thumann, W., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (12-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26226
 Center clone name: 2013_P_3

* NOTE: This record contains 74 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1	697:	contig of 697 bp in length
698	797:	gap of 100 bp
798	1493:	contig of 696 bp in length
1494	1593:	gap of 100 bp
1594	2271:	contig of 678 bp in length
2272	2371:	gap of 100 bp
2372	3081:	contig of 710 bp in length
3082	3181:	gap of 100 bp
3182	3909:	contig of 728 bp in length
3910	4009:	gap of 100 bp
4010	4683:	contig of 674 bp in length
4684	4783:	gap of 100 bp
4784	5499:	contig of 716 bp in length
5500	5599:	gap of 100 bp
5600	6287:	contig of 688 bp in length
6288	6387:	gap of 100 bp
6388	7081:	contig of 694 bp in length
7082	7181:	gap of 100 bp
7182	7882:	contig of 701 bp in length
7883	7982:	gap of 100 bp
7983	8683:	contig of 701 bp in length
8684	8783:	gap of 100 bp
8784	9484:	contig of 701 bp in length
9485	9584:	gap of 100 bp
9585	10293:	contig of 709 bp in length
10294	10393:	gap of 100 bp
10394	11087:	contig of 694 bp in length
11088	11187:	gap of 100 bp
11188	11907:	contig of 720 bp in length
11908	12007:	gap of 100 bp
12008	12592:	contig of 685 bp in length
12593	12792:	gap of 100 bp
12793	13489:	contig of 697 bp in length
13490	13589:	gap of 100 bp

13590	14301:	contig of 712 bp in length
14302	14401:	gap of 100 bp
14402	15112:	contig of 711 bp in length
15113	15212:	gap of 100 bp
15213	15896:	contig of 684 bp in length
15897	15996:	gap of 100 bp
15997	16686:	contig of 690 bp in length
16687	16786:	gap of 100 bp
16787	17491:	contig of 705 bp in length
17492	17591:	gap of 100 bp
17592	18292:	contig of 701 bp in length
18293	18392:	gap of 100 bp
18393	19107:	contig of 715 bp in length
19108	19207:	gap of 100 bp
19208	19927:	contig of 720 bp in length
19928	20027:	gap of 100 bp
20028	20720:	contig of 693 bp in length
20721	20820:	gap of 100 bp
20821	21523:	contig of 703 bp in length
21524	21623:	gap of 100 bp
21624	22335:	contig of 712 bp in length
22336	22436:	gap of 100 bp
22437	23116:	contig of 681 bp in length
23117	23216:	gap of 100 bp
23217	23911:	contig of 695 bp in length
23912	24011:	gap of 100 bp
24012	24706:	contig of 695 bp in length
24707	24806:	gap of 100 bp
24807	25484:	contig of 678 bp in length
25485	25584:	gap of 100 bp
25585	26268:	contig of 684 bp in length
26269	26368:	gap of 100 bp
26369	27079:	contig of 711 bp in length
27080	27179:	gap of 100 bp
27180	27865:	contig of 686 bp in length
27866	27965:	gap of 100 bp
27966	28667:	contig of 702 bp in length
28668	28767:	gap of 100 bp
28768	29457:	contig of 690 bp in length
29458	29557:	gap of 100 bp
29558	30259:	contig of 702 bp in length
30260	30359:	gap of 100 bp
30360	31028:	contig of 669 bp in length
31029	31128:	gap of 100 bp
31129	31827:	contig of 699 bp in length
31828	31927:	gap of 100 bp
31928	32647:	contig of 720 bp in length
32648	32747:	gap of 100 bp
32748	33458:	contig of 711 bp in length
33459	33558:	gap of 100 bp
33559	34257:	contig of 699 bp in length
34258	34357:	gap of 100 bp
34358	35045:	contig of 688 bp in length
35046	35145:	gap of 100 bp
35146	35869:	contig of 724 bp in length
35870	35969:	gap of 100 bp
35970	36652:	contig of 683 bp in length
36653	36752:	gap of 100 bp
36753	37449:	contig of 697 bp in length
37450	37549:	gap of 100 bp
37550	38233:	contig of 684 bp in length
38234	38333:	gap of 100 bp
38334	39050:	contig of 717 bp in length
39051	39150:	gap of 100 bp
39151	39831:	contig of 681 bp in length
39832	39931:	gap of 100 bp
39932	40610:	contig of 679 bp in length
40611	40710:	gap of 100 bp
40711	41415:	contig of 705 bp in length
41416	41515:	gap of 100 bp
41516	42193:	contig of 678 bp in length
42194	42293:	gap of 100 bp
42294	42984:	contig of 691 bp in length

TITLE
JOURNAL

COMMENT

* 42985 43084: gap of 100 bp
 * 43085 43776: contig of 692 bp in length
 * 43777 43876: gap of 100 bp
 * 43878 44582: contig of 705 bp in length
 * 44583 44582: gap of 100 bp
 * 44584 45387: contig of 705 bp in length
 * 45388 46202: contig of 715 bp in length
 * 45388 46202: contig of 715 bp in length
 * 46203 46302: gap of 100 bp
 * 46303 47118: contig of 716 bp in length
 * 47019 47118: gap of 100 bp
 * 47119 47822: contig of 704 bp in length
 * 47123 47922: gap of 100 bp
 * 47923 48605: contig of 683 bp in length
 * 48606 48705: gap of 100 bp
 * 48706 49416: contig of 711 bp in length
 * 49417 49516: gap of 100 bp
 * 49517 50207: contig of 691 bp in length
 * 50208 50307: gap of 100 bp
 * 50308 50989: contig of 682 bp in length
 * 50990 51089: gap of 100 bp
 * 51090 51781: contig of 692 bp in length
 * 51782 51881: gap of 100 bp
 * 51882 52575: contig of 694 bp in length
 * 52576 52675: gap of 100 bp
 * 52676 53385: contig of 710 bp in length
 * 53386 53485: gap of 100 bp
 * 53486 54177: contig of 692 bp in length
 * 54178 54277: gap of 100 bp
 * 54278 54964: contig of 687 bp in length

Query Match 84.0%; Score 16.8; DB 2; Length 58969;
 Best Local Similarity 90.0%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CQTGTCTGTCTAGTCC 20
 |||||
 Db 15353 CQTGTCTGTCTGCAAGTCTC 15372

RESULT 37
 AC118055/c
 LOCUS
 DEFINITION Homo sapiens chromosome 15 clone CTD-2013P3 map 15, LOW-PASS
 SEQUENCE SAMPLING.
 AC118055
 AC118055.1 GI:20143784
 VERSION
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 58969)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 15, clone CTD-2013P3
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 58969)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Bouckhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collumore,A.,
 Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Katat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (12-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 Smit, A.F.A. & Green, P. (1996-1997)
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26226
 Center clone name: 2013_P_3

 * NOTE: This record contains 74 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 697: contig of 697 bp in length
 * 698 797: gap of 100 bp
 * 798 1493: contig of 696 bp in length
 * 1494 1593: gap of 100 bp
 * 1594 2271: contig of 678 bp in length
 * 2272 2371: gap of 100 bp
 * 2372 3081: contig of 710 bp in length
 * 3082 3181: gap of 100 bp
 * 3182 3909: contig of 728 bp in length
 * 3910 4009: gap of 100 bp
 * 4010 4683: contig of 674 bp in length
 * 4684 4783: gap of 100 bp
 * 4784 5499: contig of 716 bp in length
 * 5500 5599: gap of 100 bp
 * 5600 6287: contig of 688 bp in length
 * 6288 6387: gap of 100 bp
 * 6388 7081: contig of 694 bp in length
 * 7082 7181: gap of 100 bp
 * 7182 7882: contig of 701 bp in length
 * 7883 7982: gap of 100 bp
 * 7983 8683: contig of 701 bp in length
 * 8684 8783: gap of 100 bp
 * 8784 9484: contig of 701 bp in length
 * 9485 9584: gap of 100 bp
 * 9585 10293: contig of 709 bp in length
 * 10294 10393: gap of 100 bp
 * 10394 11087: contig of 694 bp in length
 * 11088 11187: gap of 100 bp
 * 11188 11907: contig of 720 bp in length
 * 11908 12007: gap of 100 bp
 * 12008 12692: contig of 685 bp in length
 * 12693 12792: gap of 100 bp
 * 12793 13489: contig of 697 bp in length
 * 13490 13589: gap of 100 bp
 * 13590 14301: contig of 712 bp in length
 * 14302 14401: gap of 100 bp
 * 14402 15112: contig of 711 bp in length
 * 15113 15212: gap of 100 bp
 * 15213 15896: contig of 684 bp in length
 * 15897 15996: gap of 100 bp
 * 15997 16686: contig of 690 bp in length

TITLE
 JOURNAL
 COMMENT

* 16787 16786: gap of 100 bp bp in length
* 16787 17491: contig of 705 bp in length
* 17492 17591: gap of 100 bp
* 17592 18292: contig of 701 bp in length
* 18293 18392: gap of 100 bp
* 18393 19107: contig of 715 bp in length
* 19108 19207: gap of 100 bp
* 19208 19207: contig of 720 bp in length
* 19228 20027: gap of 100 bp
* 20028 20720: contig of 693 bp in length
* 20721 20820: gap of 100 bp
* 20821 21523: contig of 703 bp in length
* 21524 21623: gap of 100 bp
* 21624 22335: contig of 712 bp in length
* 22336 22435: gap of 100 bp
* 22436 23116: contig of 681 bp in length
* 23117 23216: gap of 100 bp
* 23217 23911: contig of 695 bp in length
* 23912 24011: gap of 100 bp
* 24012 24706: contig of 695 bp in length
* 24707 24806: gap of 100 bp
* 24807 25484: contig of 678 bp in length
* 25485 25584: gap of 100 bp
* 25585 26268: contig of 684 bp in length
* 26269 26368: gap of 100 bp
* 26369 27079: contig of 711 bp in length
* 27080 27179: gap of 100 bp
* 27180 27865: contig of 686 bp in length
* 27866 27965: gap of 100 bp
* 27967 28667: contig of 702 bp in length
* 28668 28767: gap of 100 bp
* 28768 29457: contig of 690 bp in length
* 29458 29557: gap of 100 bp
* 29558 30259: contig of 702 bp in length
* 30260 30359: gap of 100 bp
* 30360 31028: contig of 669 bp in length
* 31029 31128: gap of 100 bp
* 31129 31827: contig of 699 bp in length
* 31828 31927: gap of 100 bp
* 31928 32647: contig of 720 bp in length
* 32648 32747: gap of 100 bp
* 32748 33458: contig of 711 bp in length
* 33459 33558: gap of 100 bp
* 33559 34257: contig of 699 bp in length
* 34258 34357: gap of 100 bp
* 34358 35045: contig of 688 bp in length
* 35046 35145: gap of 100 bp
* 35146 35869: contig of 724 bp in length
* 35870 35969: gap of 100 bp
* 35970 36522: contig of 683 bp in length
* 36523 36752: gap of 100 bp
* 36753 37449: contig of 697 bp in length
* 37450 37549: gap of 100 bp
* 37550 38233: contig of 684 bp in length
* 38234 38333: gap of 100 bp
* 38334 39050: contig of 717 bp in length
* 39051 39150: gap of 100 bp
* 39151 39831: contig of 681 bp in length
* 39832 39931: gap of 100 bp
* 39932 40610: contig of 679 bp in length
* 40611 40710: gap of 100 bp
* 40711 41415: contig of 705 bp in length
* 41416 42193: gap of 100 bp
* 42194 42293: gap of 100 bp
* 42294 42984: contig of 691 bp in length
* 42985 43085: gap of 100 bp
* 43085 43776: contig of 692 bp in length
* 43777 43876: gap of 100 bp
* 43877 44582: contig of 706 bp in length
* 44583 44682: gap of 100 bp
* 44683 45387: contig of 705 bp in length
* 45388 45487: gap of 100 bp

* 45488 46202: contig of 715 bp in length
* 46203 46302: gap of 100 bp
* 46303 47018: contig of 716 bp in length
* 47019 47118: gap of 100 bp
* 47119 47822: contig of 704 bp in length
* 47823 47922: gap of 100 bp
* 47923 48605: contig of 683 bp in length
* 48606 48705: gap of 100 bp
* 48706 49416: contig of 711 bp in length
* 49417 49516: gap of 100 bp
* 49517 50207: contig of 691 bp in length
* 50208 50307: gap of 100 bp
* 50308 50989: contig of 682 bp in length
* 50990 51089: gap of 100 bp
* 51090 51781: contig of 692 bp in length
* 51782 51881: gap of 100 bp
* 51882 52575: contig of 694 bp in length
* 52576 52675: gap of 100 bp
* 52676 53385: contig of 710 bp in length
* 53386 53485: gap of 100 bp
* 53486 54177: contig of 692 bp in length
* 54178 54277: gap of 100 bp
* 54278 54964: contig of 687 bp in length

Query Match 84.0%; Score 16.8; DB 2; Length 58969;
Best Local Similarity 90.0%; Pred. No. 5.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGTGTCTGTGCTAGTCCC 20
|||||
Db 40417 CGTGTCTGTGCAAGTCTC 40398

RESULT 38
AC004252/c

LOCUS Drosophila melanogaster (P1 DS09121 (D88)) DNA linear INV 01-MAR-1998
DEFINITION AC004252 Drosophila melanogaster (P1 DS09121 (D88)) DNA sequence, complete
sequence.

AC004252 AC001962 AC001964 AC001957 AC002935 AC001961 AC001959
AC001421 AC001954 AC001417 AC001420 AC001955 AC001418 AC001958
AC001416 AC001960 AC001953 AC001952 AC001956 AC001963
AC004252.1 GI:2920806
HTG.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 59498)

AUTHORS

Celniker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,
Humast, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,
Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.
and Kimmel, B.

Sequencing of Drosophila chromosome 2R, region 58D1-58F1

Unpublished (1997)

REFERENCE 2 (bases 1 to 59498)

AUTHORS

Celniker, S.E., Aghavani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Chew, M., Doyle, C.M., Farfan, D.E., Flanagan, J., Houston, K.A.,
Humast, S.R., Karra, K., Kearney, L., Kim, S.H., Lee, B.,
Lomotan, M.A., Mazda, P., Mok, M.S., Moshrefi, A.R., Moshrefi, M.,
Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Snir, E.,
Twomey, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.L.
and Kimmel, B.

Direct Submission

Submitted (28-FEB-1998) Berkeley Drosophila Genome Project, MS

64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,

Berkeley, CA 94720, US

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site
 (<http://fruitfly.berkeley.edu/sequence-archive.html>) or send email to drosophila.genome@lbl.gov.

Library location: 1_96.
 This F1 was assembled from the following subclones: 2_g7 (AC001962), 2_h4 (AC001964), 2_d6 (AC001957), 2_h12 (AC002935), 2_a3, 2_g12 (AC001961), 2_e3 (AC001959), 2_g6 (AC001421), 1_g10 (AC001954), 1_g4 (AC001417), 2_d8 (AC001420), 1_h7 (AC001955), 2_b2 (AC001418), 2_e10 (AC001958), 1_d2 (AC001416), 2_b3 (AC001419), 2_f1 (AC001960), 1_c3 (AC001953), 1_a3 (AC001952), 2_all (AC001956), 2_g9 (AC001963).

FEATURES

source

1..59498
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /chromosomes="2R"
 /map="58D1-58F1"
 /clone="P1 DS09121 (D88)"
 /note="Does not seem to hit D62, hits D61 on sp6 end, NOT T7 end which is indicated be the map!"

ORIGIN

Query Match 84.0%; Score 16.8; DB 3; Length 59498;
 Best Local Similarity 90.0%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20
 |||||
 Db 57708 CGTGTCTCTGCTCTGCTGCC 57689

RESULT 39

AC017870

LOCUS AC017870 88933 bp DNA linear HTG 09-DEC-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.

ACCESSION AC017870

VERSION AC017870.1 GI:6553320

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 88933)

Adams, M. and Venter, J.C.

Direct Submission

Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

This sequence was identified as CDM:10212423 by the submitter.
 For more information on this record e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1..88933
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 88933;
 Best Local Similarity 90.0%; Pred. No. 4.9e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20
 |||||
 Db 74708 CGTGTCTCTGCTCTGCTGCC 74727

RESULT 40

AC151836 3/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC151836 Accession AC151836

Fragment Name	Begin	End
AC151836_0	1	110000
AC151836_1	100001	210000
AC151836_2	200001	310000
AC151836_3	300001	410000
AC151836_4	400001	485063

Continuation (4 of 5) of AC151836 from base 300001 (AC151836 Mus musculus chromosome 14)

Query Match 84.0%; Score 16.8; DB 2; Length 110000;
 Best Local Similarity 90.0%; Pred. No. 4.7e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20
 |||||
 Db 28923 CGTGGTCTGCTAGTCTCC 28904

RESULT 41

AC151419/c

LOCUS

AC151419 113386 bp DNA linear HTG 04-SEP-2004
 DEFINITION Bos taurus clone CH240-13F11, WORKING DRAFT SEQUENCE, 12 unordered pieces.

ACCESSION AC151419

VERSION AC151419.2 GI:51889783

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 113386)

REFERENCE

AUTHORS

Muzny, D., Metzker, M., Adams, C., Agbai II, O., Allen, C., Albrechts, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltrami, R., Berarducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadore, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Diu, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaekin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B., Howell, L. T., Hulyk, S., Hume, J., Ito, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorenz, H., Lozano, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercader, C., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Muridasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokeme, O., Obregon, M., Ochi-Okorie, C., Odh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quinlan, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savary, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, B., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K., Vargo, C., Verdusco, D., Villasana, D., Virk, D., Volkov, A.,

Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Xue, S., Yan, K., Yaun, Y., Yu, P., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Direct Submission

Unpublished
2 (bases 1 to 113386)

Worley, K.C.

Direct Submission

Submitted (27-AUG-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 113386)

Worley, K.C.

Direct Submission

Submitted (04-SEP-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 4, 2004 this sequence version replaced gi:51571973.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help.tmc.edu

----- Project Information

Center project name: FAOJ

Center clone name: CH240-13F11

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 110313 bases at least Q40

Consensus quality: 110834 bases at least Q30

Consensus quality: 111456 bases at least Q20

Estimated insert size: 119536; sum-of-contigs estimation

Estimated insert size: 116681; agarose-fp estimation

Quality coverage: 3x in Q20 bases; agarose-fp estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 6636: contig of 6636 bp in length
* 6637 6736: gap of unknown length
* 6737 10947: contig of 4211 bp in length
* 10948 11048: gap of unknown length
* 11048 28138: contig of 17091 bp in length
* 28139 28238: gap of unknown length
* 28239 32060: contig of 3822 bp in length
* 32061 32160: gap of unknown length
* 32161 47757: contig of 15597 bp in length
* 47757 47857: gap of unknown length
* 47858 52314: contig of 4457 bp in length
* 52315 52414: gap of unknown length
* 52415 79094: contig of 26680 bp in length
* 79095 79194: gap of unknown length
* 79195 83988: contig of 4794 bp in length
* 83989 84088: gap of unknown length
* 84089 91962: contig of 7874 bp in length
* 91963 92062: gap of unknown length
* 92063 100259: contig of 8197 bp in length
* 100260 100359: gap of unknown length
* 100360 105935: contig of 5576 bp in length
* 105936 106035: gap of unknown length
* 106036 113386: contig of 7351 bp in length.
Location/Qualifiers

FEATURES

source

1. 113386
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-13F11"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 113386;
Best Local Similarity 90.0%; Pred. No. 4.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20

|||||
Db 56342 CGTGTGCTGTGTTAATCCC 56323

RESULT 42

AL669954

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-456018 on chromosome 11,
complete sequence.

ACCESSION AL669954

VERSION AL669954.6 GI:21531129

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 119279)

AUTHORS

Whitehead, S.

TITLE

Direct Submission

Submitted (18-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquary@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jun 21, 2002 this sequence version replaced gi:2121353.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one W13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL; Sw,

SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-456018 is

from the RPCI-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.choxi.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES

source

1. 119279
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-456018"
/clone_lib="RPCI-23"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 119279;
Best Local Similarity 90.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTGCTGTGCTAGTCCC 20

|||||

Db 86209 COTGTCTGTCTAGTCC 86228

RESULT 43
AC141138/c

LOCUS AC141138 126056 bp DNA linear HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-178N14, *** SEQUENCING IN PROGRESS
***, 60 unordered pieces.

AC141138
AC141138.3 GI:30466903
HTG: HTGS PHASE1
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 126056)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huiyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G.,
Olanpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,R., Rives,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Soza,J., Steimle,M., Strong,R., Sutton,A.,
Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,A.,
Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczky,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission
Unpublished

2 (bases 1 to 126056)
Worley,K.C.

REFERENCE JOURNAL
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 126056)
Worley,K.C.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:28913269.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KDJT
Center clone name: CH230-178N14
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 112572 bases at least Q40
Consensus quality: 119554 bases at least Q30
Consensus quality: 124684 bases at least Q20
Estimated insert size: 104198; sum-of-contigs estimation
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1055: contig of 1055 bp in length
* 1056 1155: gap of unknown length
* 1156 2369: contig of 1214 bp in length
* 2370 2469: gap of unknown length
* 2470 3476: contig of 1007 bp in length
* 3477 3576: gap of unknown length
* 3577 4932: contig of 1356 bp in length
* 4933 5032: gap of unknown length
* 5033 6748: contig of 1716 bp in length
* 6749 6849: gap of unknown length
* 6849 8053: contig of 1205 bp in length
* 8054 8153: gap of unknown length
* 8154 9614: contig of 1461 bp in length
* 9615 9715: gap of unknown length
* 9715 10926: contig of 1212 bp in length
* 10927 12460: contig of 1434 bp in length
* 12461 12560: gap of unknown length
* 12561 14028: contig of 1468 bp in length
* 14029 14128: gap of unknown length
* 14129 15485: contig of 1356 bp in length
* 15485 17253: contig of 1669 bp in length
* 17254 17353: gap of unknown length
* 17354 18799: contig of 1446 bp in length
* 18800 20239: contig of 1340 bp in length
* 20240 20339: gap of unknown length
* 20340 21494: contig of 1155 bp in length
* 21495 21594: gap of unknown length
* 21595 22660: contig of 1066 bp in length
* 22661 22761: gap of unknown length
* 22761 24315: contig of 1555 bp in length
* 24316 24415: gap of unknown length
* 24416 26021: contig of 1606 bp in length
* 26022 26121: gap of unknown length
* 26122 27257: contig of 1136 bp in length
* 27258 27357: gap of unknown length

```

* 27358 28620: contig of 1263 bp in length
* 28621 28720: gap of unknown length
* 28721 29853: contig of 1233 bp in length
* 29853 30053: gap of unknown length
* 30054 32125: contig of 2072 bp in length
* 32126 32225: gap of unknown length
* 32226 33697: contig of 1472 bp in length
* 33698 33797: gap of unknown length
* 33798 35359: contig of 1562 bp in length
* 35360 35459: gap of unknown length
* 35460 37021: contig of 1562 bp in length
* 37022 37121: gap of unknown length
* 37122 38194: contig of 1073 bp in length
* 38195 38294: gap of unknown length
* 38295 40095: contig of 1801 bp in length
* 40096 40195: gap of unknown length
* 40196 42087: contig of 1892 bp in length
* 42088 42187: gap of unknown length
* 42188 43765: contig of 1578 bp in length
* 43766 43865: gap of unknown length
* 43866 44931: contig of 1066 bp in length
* 44932 45031: gap of unknown length
* 45032 46278: contig of 1247 bp in length
* 46279 46379: gap of unknown length
* 46379 48544: contig of 2166 bp in length
* 48545 48644: gap of unknown length
* 48645 49741: contig of 1097 bp in length
* 49742 49841: gap of unknown length
* 49842 51229: contig of 1388 bp in length
* 51230 51329: gap of unknown length
* 51330 53356: contig of 2027 bp in length
* 53357 53456: gap of unknown length
* 53457 55035: contig of 1579 bp in length
* 55036 55135: gap of unknown length
* 55136 56770: gap of unknown length
* 56771 58151: contig of 1381 bp in length
* 58152 58251: gap of unknown length
* 58252 60521: contig of 2270 bp in length
* 60522 60621: gap of unknown length
* 60622 62784: contig of 2163 bp in length
* 62785 62884: gap of unknown length
* 62885 65094: contig of 2210 bp in length
* 65095 65194: gap of unknown length
* 65195 67815: contig of 2621 bp in length
* 67816 67915: gap of unknown length
* 67916 69794: contig of 1879 bp in length
* 69795 69894: gap of unknown length
* 69895 72567: contig of 2673 bp in length
* 72568 72667: gap of unknown length
* 72668 74508: contig of 1841 bp in length
* 74509 74608: gap of unknown length
* 74609 77519: contig of 2911 bp in length
* 77520 77619: gap of unknown length
* 77620 80278: contig of 2659 bp in length
* 80279 80378: gap of unknown length
* 80379 83008: contig of 2630 bp in length
* 83009 83108: gap of unknown length
* 83109 87336: contig of 4228 bp in length

```

```

Query Match      84.0%; Score 16.8; DB 2; Length 126056;
Best Local Similarity 90.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
Db 35984 CATGTCTGTGCTAGTCC 35965

RESULT 44
AC148290 132181 bp DNA linear PLN 07-MAY-2004
LOCUS Medicago truncatula chromosome 2 BAC clone mth2-45h12, complete
DEFINITION

```

```

sequence.
AC148290
VERSION AC148290.3 GI:47077943
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 132181)
AUTHORS Town, C.D., Tallon, L.J., Arbogast, T., Althoff, R., Hine, E.,
Monaghan, E., Smith, S.A., Utterback, T., Feidblyum, T., Koo, H. and
Cheung, F.
TITLE Medicago truncatula BAC genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 132181)
AUTHORS Town, C.D.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 132181)
AUTHORS Town, C.D.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 132181)
AUTHORS Town, C.D.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT On May 7, 2004 this sequence version replaced gi:45120164.
mononucleotide (T) repeat of 21 to 22 nucleotides.
FEATURES
source
1. .132181
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/chromosome="2"
/clone="mth2-45h12"

ORIGIN
Query Match      84.0%; Score 16.8; DB 8; Length 132181;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTGTGCTAGTCC 20
Db 47236 CGTGTCTGTGCTAGTCC 47255

RESULT 45
AC151300 134430 bp DNA linear HTG 25-AUG-2004
LOCUS Xenopus tropicalis clone CH216-107C14, WORKING DRAFT SEQUENCE, 2
DEFINITION
Xenopus tropicalis
Xenopus tropicalis (Silurana tropicalis)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 134430)
AUTHORS DOE Joint Genome Institute.
TITLE Unpublished
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 134430)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2004) Production Genomics Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA

```

```

94598-1698, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 3608807
Center clone name: CHORI-216_107C14
-----
Summary Statistics
Consensus quality: 133976 bases at least Q40
Consensus quality: 134224 bases at least Q30
Consensus quality: 134293 bases at least Q20
Estimated insert size: 133000; agarose-fp estimation
Estimated insert size: 134330; sum-of-contigs estimation
Quality coverage: 11.15 in Q20 bases; agarose-fp estimation
Quality coverage: 11.04 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 50955: contig of 50955 bp in length
* 50956 51055: gap of unknown length
* 51056 134430: contig of 83375 bp in length.
* 134430
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="CH216-107C14"
/clone_lib="CHORI-216 Xenopus tropicalis BAC library"

ORIGIN
Query Match      84.0%; Score 16.8; DB 2; Length 134430;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGTGTCTGCTGCTAGTCCC 20
| | | | | | | | | | | | | | | | | | | |
Db 62193 CATGTGCTCTGCTAGTCCC 62212

RESULT 46
AC145410/c
LOCUS          147434 bp      DNA      linear      HTG 06-SEP-2003
DEFINITION    Bos taurus clone RP42-227G24, WORKING DRAFT SEQUENCE, 3 ordered
               pieces
ACCESSION     AC145410
VERSION       AC145410.2 GI:34494894
KEYWORDS      HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE        Bos taurus (cow)
ORGANISM      Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 147434)
Anttonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghigbi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masfello,C.,
Maskeri,B., McDowell,J., Paquirigan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddix-Duque,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 147434)
Green,E.D.
Direct Submission
Submitted (02-JUL-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 147434)
Green,E.D.
Direct Submission
Submitted (06-SEP-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Sep 6, 2003 this sequence version replaced gi:32401574.
-----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
-----
Project Information
Center project name: eyb
Center clone name: 227G24

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 147115 bases at least Q40
Consensus quality: 147204 bases at least Q30
Consensus quality: 147222 bases at least Q20
Insert size: 145000; agarose-fp
Insert size: 147234; sum-of-contigs
Quality coverage: 12.5ix in Q20 bases; agarose-fp
Quality coverage: 12.32x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11857: contig of 11857 bp in length
* 11858 11957: gap of unknown length
* 11958 140052: contig of 128095 bp in length
* 140053 140152: gap of unknown length
* 140153 147434: contig of 7282 bp in length.
Location/Qualifiers
1. 147434
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-227G24"
/clone_lib="RP42"
1. 11857
/note="assembly_fragment"
clone end:SP6
vector side:left
11958..140052
/note="assembly_fragment"
133983..147434
/note="clone overlaps with GenBank Accession Number

FEATURES
source
misc_feature
misc_feature
misc_feature

```

```

misc_feature      AC145443 clone RP42-38D1 (center project name eyc)
                  140153 ..147434
                  /note="assembly_fragment
                  clone_end:T7
                  vector_side:right"

ORIGIN
Query Match      84.0%; Score 16.8; DB 2; Length 147434;
Best Local Similarity 90.0%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CGTGTGCTGCTAGTCC 20
Db      48782 CGTGTGCTGCTAGTCC 48763

RESULT 47
AB020868
LOCUS      AB020868      153472 bp      DNA      linear      PRI 21-MAY-1999
DEFINITION Homo sapiens genomic DNA of 8p21.3-p22 anti-oncogene of
             hepatocellular colorectal and non-small cell lung cancer , segment
             11/11.
ACCESSION      AB020868
VERSION      AB020868.1 GI:4003388
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      1 (sites)
              Nakamura,Y. and Isomura,M.
TITLE      Homo sapiens 1,210,381bp genomic DNA of 8p21.3-p22 anti-oncogene of
             hepatocellular colorectal and non-small cell lung cancer region
             2 (bases 1 to 153472)
JOURNAL      Hirakawa,M., Yamaguchi,H., Imai,K. and Shimada,J.
AUTHORS      Direct Submission
TITLE      Submitted (30-NOV-1998) Mita Hirakawa, Japan Science and Technology
             Corporation (JST), Advanced Database Department; 5-3, Yonban-cho,
             Chiyoda-ku, Tokyo 102-0028, Japan (E-mail:mika@tokyo.jst.go.jp,
             Tel:81-3-5214-8491, Fax:81-3-5214-8470)
JOURNAL      This sequence is conducted by Japanese Foundation for Cancer
             Research as a JST sequencing team
COMMENT      Principal Investigator: Yusuke Nakamura Ph.D.
             Phone:+81-3-5449-5372, Fax:+81-3-5449-5433,
             yusuke@hgc.ims.u-tokyo.ac.jp
             The sequence is submitted by:Human Genome Sequencing in ALIS
             project of JST
             Japan Science and Technology Corporation (JST)
             5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081 Japan
             For further information about this sequence, including its location
             and relationship to other sequences, please visit our sequence
             archive Web site
             (http://www-alis.tokyo.jst.go.jp/HGS/top.html)
             or send email to webmaster@www-alis.tokyo.jst.go.jp.

FEATURES             source
             1..153472
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="8"
                /map="8p21.3-p22"
             7141..7378
                /note="atSG1569;The location is between each flanking site
                of PCR primers."
                /db_xref="GDB:4564419"

STS
Query Match      84.0%; Score 16.8; DB 9; Length 153472;
Best Local Similarity 90.0%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
QY      1 CGTGTGCTGCTAGTCC 20

```

```

Db      16465 CCTGTGCTGCTATTCCTCC 16484

RESULT 48
AC116903
LOCUS      AC116903      159893 bp      DNA      linear      PRI 06-AUG-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-152L20, complete sequence.
ACCESSION      AC116903
VERSION      AC116903.3 GI:22004395
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      1 (bases 1 to 159893)
              Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
              Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
              Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collumore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
              Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
              Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,N., Jones,C.,
              Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
              Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
              MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
              McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L.,
              Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
              Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
              Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
              Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
              Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
              Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
              Strauss,N., Subramanian,A., Talamas,J.J., Tesfaye,S., Theodore,J.,
              Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
              Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE      Direct Submission
JOURNAL      Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS      3 (bases 1 to 159893)
              Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
              Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
              Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collumore,A.,
              Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
              Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
              Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
              Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
              Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
              Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
              Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrim,J.,
              Meneus,L., Mihoa,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
              Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
              O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
              Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
              Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
              Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
              Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
              Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
              Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE      Direct Submission
JOURNAL      Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS      4 (bases 1 to 159893)
              Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

```


Contact: sequence submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26182
 Center clone name: 268_L_17

FEATURES source Location/Qualifiers
 1..160580
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-268L17"
 /clone_lib="RPC1-11 Human Male BAC"
 1558..1589
 /rpt_family="CGG)n"
 1627..1680
 /rpt_family="C-rich"
 2078..2110
 /rpt_family="G-rich"
 complement(3581..3697)
 /rpt_family="L2"
 complement(5081..5216)
 /rpt_family="MIR"
 6033..6061
 /rpt_family="(TTTG)n"
 complement(6422..6548)
 /rpt_family="L2"
 complement(6549..6652)
 /rpt_family="MIR"
 6791..7072
 /rpt_family="MER63"
 complement(7460..7929)
 /rpt_family="MLT1J"
 8499..8703
 /rpt_family="MIR"
 8750..8773
 /rpt_family="AT-rich"
 8774..8887
 /rpt_family="LIME4A"
 9095..9145
 /rpt_family="L2"
 13294..14510
 /rpt_family="HSMAR2"
 14515..14696
 /rpt_family="(TTC)n"
 14697..14765
 /rpt_family="(TCC)n"
 complement(14768..14900)
 /rpt_family="FLAM_A"
 14905..14954
 /rpt_family="HSMAR2"
 15070..15158

Query Match 84.0%; Score 16.8; DB 9; Length 160580;
 Best Local Similarity 90.0%; Pred. No. 4.4e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20
 |||||
 DB 128514 CGTGTCTCTGCAAGTCTC 128533

RESULT 50
 AP006203
 LOCUS AP006203 168921 bp DNA linear PRI 18-FEB-2003
 DEFINITION Homo sapiens genomic DNA, chromosome 8, clone:RP11-110N16, complete sequence.
 ACCESSION AP006203
 VERSION AP006203.1 GI:28411639
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens genomic DNA
 Published Only in Database (2003)
 2 (bases 1 to 168921)
 Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
 Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (14-FEB-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel.81-45-503-9111, Fax:81-45-503-9170)
 FEATURES source Location/Qualifiers
 1..168921
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /clone="RP11-110N16"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 168921;
 Best Local Similarity 90.0%; Pred. No. 4.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGTGTCTCTGCTAGTCCC 20
 |||||
 DB 73083 CGTGTCTCTGCTATGCC 73102

Search completed: March 15, 2005, 19:01:07
 Job time : 1721 secs

THIS PAGE BLANK (USPTO)